

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 04-12-04
Searcher: Barbara C 2528
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 25
Number of Searches: _____
Number of Databases: 1

Search Site	Vendors
<input type="checkbox"/> STIC	<input type="checkbox"/> IG
<input type="checkbox"/> CM-1	<input type="checkbox"/> STN
<input type="checkbox"/> Pre-S	<input type="checkbox"/> Dialog
Type of Search	
<input type="checkbox"/> N.A. Sequence	<input type="checkbox"/> APS
<input type="checkbox"/> A.A. Sequence	<input type="checkbox"/> Geninfo
<input type="checkbox"/> Structure	<input type="checkbox"/> SDC
<input type="checkbox"/> Bibliographic	<input type="checkbox"/> DARC/Questel
	<input checked="" type="checkbox"/> Other

This page Blank (uspto)

STIC-Biotech/ChemLib

118960

From: Leffers, Gerald
Sent: Wednesday, April 07, 2004 1:37 PM
To: STIC-Biotech/ChemLib
Subject: 09/430,590

Please interference search/search SEQ ID NO: 3 (~ 6.5 kb). Thank you. Gerry Leffers

Gerald G. Leffers Jr., PhD
Primary Examiner, Art Unit 1636
Remsen Building, Room 02A69
(571) 272-0772

118960
APR 7 2004
STIC-BIOTECH/CHM LIB
SEARCHED INDEXED
SERIALIZED FILED

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

This page blank (uspto)

Db	901	TTGCTGCAGGGGAGAGAACTCGAGGAATCTCACACACTCACACGTGTGCA	960	QY	2041	CCCATATGGAGTGAACATGCTACCAACTGGAGATAAGACGATTTAATTCCA	2100
QY	961	TGAGTAGTGTGCTGAGCTCTCAAATCTCAAGGTCAAGCTGACTCAATATGCA	1020	Do	2041	CCCATATGGAGTGAACATGCTACCAACTGGAGATAAGACGATTTAATTCCA	2100
Db	961	TGAGTAGTGTGCTGAGCTCTCAAATCTCAAGGTCAAGCTGACTCAATATGCA	1020	QY	2101	TTGATGCTCAATCATATGTCATTAGAAATCTGTGTACAAATACCGGCT	2160
QY	1021	GAATCAGAGCTTAAAGTTAGTTGTTGAAATTCATGCTCATACTCAAGTGT	1080	Do	2101	TTGATGCTCAATCATATGTCATTAGAAATCTGTGTACAAATACCGGCT	2160
Db	1021	GAATCAGAGCTTAAAGTTAGTTGTTGAAATTCATGCTCATACTCAAGTGT	1080	QY	2161	CGTACTTCACCTCAAAAGAGAGTCCTAAAGATGCTGCAATGAAAGAGA	2220
QY	1081	TCAAGTGAATATGATCATGATCATGACTCCCAAGTGTCTTGAATTCATGCT	1140	Do	2161	CGTACTTCACCTCAAAAGAGAGTCCTAAAGATGCTGCAATGAAAGAGA	2220
Db	1081	TCAAGTGAATATGATCATGATCATGACTCCCAAGTGTCTTGAATTCATGCT	1140	QY	2221	ATCGAATCCAAACAGAGAGTCACATCATTCAGAAAGAAAGCCTGAGAGA	2280
QY	1141	TGATCATCCGATATGCTAGATCTCCAGTCATGAACTACAAATATGACTAAT	1200	Do	2221	ATCGAATCCAAACAGAGAGTCACATCATTCAGAAAGAAAGCCTGAGAGA	2280
Db	1141	TGATCATCCGATATGCTAGATCTCCAGTCATGAACTACAAATATGACTAAT	1200	QY	2281	TGAGAGCTCATGCTACTCTGGCCATTAGTCGGAAATACAGGTATT	2340
QY	1201	TATGTTAGTTTACAGCACCGAGAGAACCAAGAACAGACTGCTCCTGCA	1260	Do	2281	TGAGAGCTCATGCTACTCTGGCCATTAGTCGGAAATACAGGTATT	2340
Db	1201	TATGTTAGTTTACAGCACCGAGAGAACCAAGAACAGACTGCTCCTGCA	1260	QY	2341	AACTCTTTATAGATGACACATACGGTTACATTGAGGATTAACTAAAGCAGA	2400
QY	1261	ACAACTCAAAAGAGAACTGAATCANGAAGGAATAGAACATCCAAATCGA	1320	Do	2341	AACTCTTTATAGATGACACATACGGTTACATTGAGGATTAACTAAAGCAGA	2400
Db	1261	ACAACTCAAAAGAGAACTGAATCANGAAGGAACCAAGAGAACACTCACTG	1320	QY	2401	GGTAAGAGTCCTTATTCACGATTAACATCGGTTACATTGAGGATTAACTAAAGCAGA	2460
QY	1321	TAACGATAAAGTGAAGAAGAAAGAAAGAAACACTACTGGATGAAACAGG	1380	Do	2401	GGTAAGAGTCCTTATTCACGATTAACATCGGTTACATTGAGGATTAACTAAAGCAGA	2460
Db	1321	TAACGATAAAGTGAAGAAGAAAGAAACACTACTGGATGAAACAGG	1380	QY	2461	GGCATACTTCAGAGTCATAGCTCTGGCCATTAGTCGGCT	2520
QY	1381	TGCTGTTCTTAAATGTTGAAATGCAAACTCACTGCTTCACTGAACT	1440	Do	2461	GGCATACTTCAGAGTCATAGCTCTGGCCATTAGTCGGCT	2520
Db	1381	TGCTGTTCTTAAATGTTGAAATGCAAACTCACTGCTTCACTGAACT	1440	QY	2521	CGTATTGGAGGAGTCATAGCCATCTGGCTGAGTTATGGCTCGGAGGT	2580
QY	1441	AGAAATTCTCACTCTTAAATGCTCTTCACTGAACTATGAACTAAGTTAGGT	1500	Do	2521	CGTATTGGAGGAGTCATAGCCATCTGGCTGAGTTATGGCTCGGAGGT	2580
Db	1441	AGAAATTCTCACTCTTAAATGCTCTTCACTGAACTATGAACTAAGTTAGGT	1500	QY	2581	TGTTAAATGATTACAGTTACAGATTAAGGATGTTGAACTGTGTCACAA	2640
QY	1501	TAACAGTATAGTGTGATGATCTGGCCACATATGTTAGGT	1560	Do	2581	TGTTAAATGATTACAGTTACAGATTAAGGATGTTGAACTGTGTCACAA	2640
Db	1501	TAACAGTATAGTGTGATGATCTGGCCACATATGTTAGGT	1560	QY	2641	ACTCAAGTGTGATTTATGTTGATCATTACAGTTACAGATTAAGGAGGAA	2760
QY	1561	ATTGCTGAATTTAAGGACCAACTATGGTTCTGCTGCTGATGCTCATTGA	1620	Do	2641	ACTCAAGTGTGATTTATGTTGATCATTACAGTTACAGATTAAGGAGGAA	2760
Db	1561	ATTGCTGAATTTAAGGACCAACTATGGTTCTGCTGCTGATGCTCATTGA	1620	QY	2701	TCTTCATCAAGGGCAACCCCTATGGTGTCTTATCAATTAAGGAGGAA	2760
QY	1621	TAACAGTGTGATGTTGATGATTAATCAGAGTCGATGTTGAAACAAATTGAGA	1680	Do	2701	TCTTCATCAAGGGCAACCCCTATGGTGTCTTATCAATTAAGGAGGAA	2760
Db	1621	TAACAGTGTGATGTTGATGATTAATCAGAGTCGATGTTGAAACAAATTGAGA	1680	QY	2761	CTACCGTTCTCTTTCGATGTTGCTGTTATGAACTGCACTGAAAGAA	2820
QY	1681	TAACAGTGTGATGTTGATGATTAATCAGAGTCGATGTTGAAACAAATTGAGA	1740	Do	2761	CTACCGTTCTCTTTCGATGTTGCTGTTATGAACTGCACTGAAAGAA	2820
Db	1681	TAACAGTGTGATGTTGATGATTAATCAGAGTCGATGTTGAAACAAATTGAGA	1740	QY	2821	CCCTAACGGTTACATCAACTAAAGGCTCTTCATGGCATCGATGGCTGTGATGG	2880
QY	1741	AGGATTAATGTTCTTAAATGAACTGTTGATGATGTTGCTGATGGCT	1800	Do	2821	CCGTTAGGGAGTACATCAACTAAAGGCTCTTCATGGCATCGATGGCTGTGATGG	2880
Db	1741	AGGATTAATGTTCTTAAATGAACTGTTGATGATGTTGCTGATGGCT	1800	QY	2881	CTACGGCTGAGTGTGTTAGTATTAGGTTGCTAAAGATGCGGTGATATT	2940
QY	1801	TACTATATTCTTCAAGGAGAATCTGCTGATTTGTTGAGGATGTTGCTCAGTCA	1860	Do	2881	CTACGGCTGAGTGTGTTAGTATTAGGTTGCTAAAGATGCGGTGATATT	2940
Db	1801	TACTATATTCTTCAAGGAGAATCTGCTGATTTGTTGAGGATGTTGCTCAGTCA	1860	QY	2941	CTTAGCCCTAAATGCGTAAATGCGTAAATGCGTAAATGCGTAAATGCGTAA	3000
QY	1861	AGAATTTGATGATGTTGATGATGTTGCTCAGTCAATGCTGATGGCT	1920	Do	3001	CTTATCCCTACCTCTGGAGGTATGAGGTTAACTCTCTCTCTCAAA	3060
Db	1861	AGAATTTGATGATGTTGATGATGTTGCTCAGTCAATGCTGATGGCT	1920	QY	3001	CTTATCCCTACCTCTGGAGGTATGAGGTTAACTCTCTCTCTCAAA	3060
QY	1921	CCAGATGACAGATAATCAAGTATGAAATGTCAGATCAAGACATGATTA	1980	Do	3061	ACTGGGGCTCTAGGAGGATGCGGAAACATATGCGGAAGTGTGAAATCA	3120
Db	1921	CCAGATGACAGATAATCAAGTATGAAATGTCAGATCAAGACATGATTA	1980	QY	3061	ACTGGGGCTCTAGGAGGATGCGGAAACATATGCGGAAGTGTGAAATCA	3120
QY	1981	TAGTTCTGAGCTTAAATCTTGAAGGAGGTGTTGTTAGTGTGAAATTG	2040	Do	3061	ACTGGGGCTCTAGGAGGATGCGGAAACATATGCGGAAGTGTGAAATCA	3120
Db	1981	TAGTTCTGAGCTTAAATCTTGAAGGAGGTGTTGTTAGTGTGAAATTC	2040	QY	3061	ACTGGGGCTCTAGGAGGATGCGGAAACATATGCGGAAGTGTGAAATCA	3120

Db	5281	TGACAATATGAAAGGTTGCCATAATTCGTGAAAGTATTCCGAGATGATTG	5340
Qy	5341	AAAGGTCGAACTGAAACATGCTTCACCCGTATGTTAAACATATACCAAGTGT	5400
Db	5341	AAAGGTTCGAAATGAAACATGCTTCACCCGTATGTTAAACATATACCAAGTGT	5400
Qy	5401	TGCCCTTTATGGCTGCCAACACACGTCGATATCGCTCTAGGATGTTG	5460
Db	5401	TGCCCTCTTATGGCTGCCAACACACGTCGATATCGCTCTAGGATGTTG	5460
Qy	5461	GTGCGGTCTAACTCCAAATGTCGATGTTGAGATGAGATGAGATGTTG	5520
Db	5461	GTGCGGTCTAACTCCAAATGTCGATGTTGAGATGAGATGAGATGTTG	5520
Qy	5521	TAGTATTCAAATAGCATGGATACATGGTACATGCAAAAGAACGATGAT	5580
Db	5521	TAGTATTCAAATAGCATGGTACATGCAAAAGAACGATGAT	5580
Qy	5581	ACACCAAATATTGTTACGAACTGCTGCAACACAGTCAGATATCGCT	5640
Db	5581	ACACCAAATATTGTTACGAACTGCTGCAACACAGTCAGATATCGCT	5640
Qy	5641	TAGAAATCTATAGGAACTTGATTATGTTGAAATTTGGCTCAATGGGAC	5700
Db	5641	TAGAAATCTATAGGAACTTGATTATGTTGAAATTTGGCTCAATGGGAC	5700
Qy	5701	CAAAACAAACGCTCATGCAAACTGCTGCAACAGCTGACGATGTTG	5760
Db	5701	CAAAACAAACGCTCATGCAAACTGCTGCAACAGCTGACGATGTTG	5760
Qy	5761	TACATGTTGAAAGCTATGAAATAAACATTAACTGAAATTGGTTGAGT	5820
Db	5761	TACATGTTGAAAGCTATGAAATAAACATTAACTGAAATTGGTTGAGT	5820
Qy	5821	TAGATACATGGCATAGAACGACAGCTGTTGATAAGTTGAAATATTG	5880
Db	5821	TAGATACATGGCATAGAACGACAGCTGTTGATAAGTTGAAATATTG	5880
Qy	5881	TCACCCATGCAACATAGATCTCTATAGTTCTAGCCATTGTCAGATA	5940
Db	5881	TCACCCATGCAACATAGATCTCTATAGTTCTAGCCATTGTCAGATA	5940
Qy	5941	AGTATTGAACTCTATGAGAACAAATAATTACCGGATGTTGATGACTG	6000
Db	5941	AGTATTGAACTCTATGAGAACAAATAATTACCGGATGTTGATGACTG	6000
Qy	6001	TCTAAGTGTGCTAAATCAAGCATCTGTTGAGGTTATGAACTG	6060
Db	6001	TCTAAGTGTGCTAAATCAAGCATCTGTTGAGGTTATGAACTG	6060
Qy	6061	AGATATAAACACTGATACAAATGAAAGGGAGAAGTGGATTATCATT	6120
Db	6061	AGATATAAACACTGATACAAATGAAAGGGAGAAGTGGATTATCATT	6120
Qy	6121	ATATTGTTATGTCATCAGGGGAGTTGGCTATTGCACTGTTGTCAGACTG	6180
Db	6121	ATATTGTTATGTCATCAGGGGAGTTGGCTATTGCACTGTTGTCAGACTG	6180
Qy	6181	ATCATGAAATGGTTATGAGAACATTTCATACACTCAGGGT	6240
Db	6181	ATCATGAAATGGTTATGAGAACATTTCATACACTCAGGGT	6240
Qy	6241	GACAGAACTATATGTTGAGATAATAGGGTATGAAATCCGAGA	6300
Db	6241	GACAGAACTATATGTTGAGATAATAGGGTATGAAATCCGAGA	6300
Qy	6301	TATCACCGAGATGAGGGAGTTGCAATATATCTTGAAATTAATCTG	6360
Db	6301	TATCACCGAGATGAGGGAGTTGCAATATATCTTGAAATTAATCTG	6360
Qy	6361	AATTCGACTATACGAACTAGAGCTGTCACGTCAGCTGAAAGTTTATT	6420
Db	6361	AATTCGACTATACGAACTAGAGCTGTCACGTCAGCTGAAAGTTTATT	6420
Qy	6421	CCATCA	6426
Db	6421	CCATCA	6426

/db xref="GI:3273718" translation="KTGASINCWNCNSKTPWENSSINASVWAFGRFN KLVIDGTTISVNNKLNISTVADGTLADCGDLIRVSTLE NTULPESSFNMLVSIKQIBERGVNLTRVEVNQVNAPIASRNALDYMCPQ FSESSLECFYQGLDMSNAMDKSSMSEYQEDYSSRLNIDTEVYL DVE-SPYQEQLPQDMLNEMHNSKMEYQEDYSSRLNIDTEVYL DCKVCLSNKQREHNNHSEKRSRERHLADTIGPFSENKWLTSDEHGTI EGTITDKRDKVBLQRIKINRNHRPDKVAYRSNAPFOPSDLAEGFWRETIA YSPLENLGAEVNKLLQIYR-VVTLGPOQILLYVQISMTMHTPGRSKQ PIGCYQLSEGFNRPFALDCTVTSNAEKRGITSTKQFSSMAGVYGSDC FSYVILKMRNCDTILSPVNTIRSYEVINSVNLNTPMHLVPAEGQGRSGA QY 421 TAACGAGGGAGGTATGSAAGGTGTGTCAGCTAAGCTATAGTAGGTGATG 540 421 TAACGAGGGAGGTATGSAAGGTGTGTCAGCTAAGCTATAGTAGGTGATG 686 627 TAACGAGGGAGGTATGSAAGGTGTGTCAGCTAAGCTATAGTAGGTGATG 687 ACATATCAAGGTGATCAATATGCTTCAATAATTATGACTACCTAGTTGGC 746 541 GCTCGTATCGAAGCGGAAATGGAATAGGAATCAGGATTAAATTCCAGGTG 600 747 CTCGCGTAAATGAAAGCGTGTAAATGAGTAAATTCAGTTCCAGTTG 805 807 TAACCGATGTTGCAATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 866 601 TAACCGATGTTGCAATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 660 661 AAAGGTGAGGTATTATTATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 925 867 AAAGGTGAGGTATTATTATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 721 TCAATAGTCGATTATCATGATTCAGTGTCTAAGCTACAAAGCACTGCAAAAGGAA 780 927 TCAATAGTCGATTATCATGATTCAGTGTCTAAGCTACAAAGCACTGCAAAAGGAA 985 QY 781 CTTAAACAAACTCTATGATGCTTCAATGATTAAGATTACCAATTAGTAA 840 Db 987 CTTAAACAAACTCTATGATGCTTCAATGATTAAGATTACCAATTAGTAA 1046 Db QY 841 GATCTCTGAGTTGTAACGTTAACATGAGTTAGCTGTTCTATACCTCATA 900 Db 1047 GATCTCTGAGTTGTAACGTTAACATGAGTTAGCTGTTCTATACCTCATA 1105 Db QY 901 TTGCTCAGGGAAAGAAACTCTAGGAAATCTACCTACACACTCAACGTGTG 960 Db 1107 TTGCTCAGGGAAAGAAACTCTAGGAAATCTACCTACACACTCAACGTGTG 1166 Db QY 961 TCAATGTCCTAGTCTCAAATOTCATAGTTAGTGTCTGTACTCAATCTGAA 1020 Db 1167 TCAATGTCCTAGTCTCAAATCTACCTACAGTGTCTGTACTCAATCTGAA 1226 Db QY 1021 GAAATCAGAGGTTTAAAGTGTGTTGATATTCTGCTCATACTACTCAAGTGGAT 1080 Db 1227 GAAATCAGAGGTTTAAAGTGTGTTGATATTCTGCTCATACTCAAGTGGAT 1286 Db QY 1081 TCAAGCTGACATGATCACATGAGTCTCCAGTGTCTCTACCAAGCTGAGAAATG 1140 Db 1287 TCAAGCTGACATGATCACATGAGTCTCCAGTGTCTCTACCAAGCTGAGAAATG 1346 Db QY 1141 TGATCTCTGATTAGTCTAGTGTGACATCCACCAACAAATGACTATCT 1200 Db 1347 TGATCTCTGATTAGTCTAGTGTGACATCCACCAACAAATGACTATCT 1405 Db QY 1201 TATTTAGTTTACCCAGGAGAACCAAAAGAACAGCAGAACACTCT 1280 Db 1407 TATTTAGTTTACCCAGGAGAACCAAAAGAACAGCAGAACACTCT 1466 Db QY 1261 AACATCTCAAAGAGAACCTGAAATCAAGAACAGAACATCCAAATCAGA 1320 Db 1467 AACATCTCAAAGAGAACCTGAAATCAAGAACAGAACATCCAAATCAGA 1526 Db QY 1321 TCACTAAAGTGAAGAAAGAAAGAAAGAAACTCTACGTGAAACACAGG 1380 Db 1527 TCACTAAAGTGAAGAAAGAAAGAAACTCTACGTGAAACACAGG 1586 Db QY 1381 TCTGCTCTATAATGTTGATGATCATATTGCAAGAACAGCTTCAGT 1440 Db 1587 TCTGCTCTATAATGTTGATGATCATATTGCAAGAACACAGTTCAGT 1646

QY	1441	AGAAATTCTTATTCTTAATGCTCTTGGACCTAATGATTAAAGGTTAAGGT	1500	Db	2727		2.77
Db	1647	AGAAATTCTTATTCTTAATGCTCTTGGACCTAATGATTAAAGGTTAAGGT	1706	QY	2781	TGTTAAATGATTACAGATTACAGATTCAGGATCGTGTGACACTTGCTCACAA	2786
QY	1501	TRACAGTATTTAGTGTATGTTACCTGTCGACCAATATCTGTTGAAACAAAGAT	1560	Db	2717	TGTTAAATGATTACAGATTACAGATTCAGGATCGTGTGACACTTGCTCACAA	2846
Db	1707	TRACAGTATTTAGTGTATGTTACCTGTCGACCAATATCTGTTGAAACAAAGAT	1766	QY	2641	ACTCAAGTGTATTATGTTACATATCTGTTACATGAACTACCTCA	2700
QY	1561	ATTCGAAATTGAGCTTAAAGGACGACATGGAGTTCTGCTGTTGCTCATTGA	1520	Db	2847	ACTCAAGTGTATTATGTTACATATCTGTTACATGAACTACCTCA	2906
Db	1767	ATTCGAAATTGAGCTTAAAGGACGACATGGAGTTCTGCTGTTGCTCATTGA	1826	QY	2701	TCTTCACTCAGGAAACCCATTGGTGTCTATTACATTAAGTGTGAGGAAATT	2760
QY	1621	AGCAGATGTTGTTGGTGTCTTACATTACAGGTCGCGATATGCTCGATTGA	1680	Db	2907	TCTTCACTCAGGAAACCCATTGGTGTCTATTACATTAAGTGTGAGGAAATT	2966
Db	1827	AGCAGATGTTGTTGGTGTCTTACATTACAGGTCGCGATATGCTCGATTGA	1886	QY	2711	CTACCGTTCTTCCATCGATGTTGCTCATTTAGTAAGTGTGAGGAAAGAA	2820
QY	1681	TACATGTTTACCAAGAAACTCCTTAATCTGAGTTGAAACAAATGAGAACG	1740	Db	2967	CTACCGTTCTTCCATCGATGTTGCTCATTTAGTAAGTGTGAGGAAAGAA	3026
Db	1887	TACATGTTTACCAAGAAACTCCTTAATCTGAGTTGAAACAAATGAGAACG	1946	QY	2821	CCGTTACGGAGTACATCACTAAAGGCTCTCATCGATCATGGTGTGATGG	2880
QY	1741	AGGATTAATGTTCTTAACTTAAGATGTTGTTTACCAAATGAGCTCC	1800	Db	3027	CCGTTACGGAGTACATCACTAAAGGCTCTCATCGATCATGGTGTGATGG	3086
Db	1947	AGGATTAATGTTCTTAACTTAAGATGTTGTTTACCAAATGAGCTCC	2006	QY	2881	CTACGCTACGAGTTGTTAGTATAGTGTGTTGCTAAAGGATATCGGGTCTG	2940
QY	1801	TACTATTTATGTTCTTCAAGGAGAATGCTGGATCTTATGGCTTCAATCAGTA	1860	Db	3087	CTACGCTACGAGTTGTTAGTATAGTGTGTTGCTAAAGGATATCGGGTCTG	3146
Db	2007	TACTATTTATGTTCTTCAAGGAGAATGCTGGATCTTATGGCTTCAATCAGTA	2066	QY	2941	CCITAGCCCTAATGTCGTTATGGGAGCTATGGTTAACTCTCTCTCAA	3000
QY	1861	AGATCTTGTAGTGTATTGTTGATGGTTGGCAAGATGTTGTCATGCTA	1920	Db	3147	CCITAGCCCTAATGTCGTTATGGGAGCTATGGTTAACTCTCTCTCAA	3206
Db	2067	AGAATCTTAACTTGTGTTGTTGATGTTGATGTTGCTTTCATGCTA	2126	QY	3001	CITATCCACTACACTATGTCACATGTTCTATGGCTGAGGTTCCGGGAGGA	3060
QY	1921	CGAAGTGTACAAAGATAATGTTGACGGAGTTGATGTTAGTGTGAAATT	1980	Db	3207	CTTATCCACTACACTATGTCACATGTTCTATGGCTGAGGTTCCGGGAGGA	3266
Db	2127	CGAAGTGTACAAAGATAATGTTGACGGAGTTGATGTTAGTGTGAAATT	2186	QY	3061	ACTGGGCCCTACTACGTTGGGAGCAGATGTTGAAATGCAATCAA	3120
QY	1981	TAGTTCCTGACGACATTAAATCTTGTGACGGAGTTGATGTTAGTGTGAAATT	2040	Db	3267	ACTGGGCCCTACTACGTTGGGAGCAGATGTTGAAATGCAATCAA	3326
Db	2187	TAGTTCCTGACGACATTAAATCTTGTGACGGAGTTGATGTTAGTGTGAAATT	2246	QY	3121	TCAGGTGAGCATGCGCAAGAGGATGATAGTCAGGCGAGTCGTTACTTAC	3180
QY	2041	CCCATATGGAGTTGAAACATGTTGACGGAGATAAGCAATTATTTATC	2100	Db	3327	TCAGGTGAGCATGCGCAAGAGGATGATAGTCAGGCGAGTCGTTACTTAC	3386
Db	2247	CCCATATGGAGTTGAAACATGTTGACGGAGTTGAGATAAGCAATTATTTATC	2306	QY	3181	TAGGTTACAGTTCAAGGATGTTATGTTAGTGTGTTAGTGTGAAATGCAATCAA	3240
QY	2101	TTGAGTCATAATGTCATGAGAAATCTGTTTACAAATTACGGGCT	2160	Db	3387	TCAGGTGAGCATGCGCAAGAGGATGATAGTCAGGCGAGTCGTTACTTAC	3446
Db	2307	TTGAGTCATAATGTCATGAGAAATCTGTTTACAAATTACGGGCT	2366	QY	3241	GAATCCCGATGTTCTAACCTCTCAACTAATGAGATCACAATGATGTTACTTAC	3300
QY	2161	GTACTCTACACTCTAAAGGAGTCATCATCATGTTGTTAGTGTGTT	2220	Db	3447	GAATCCCGATGTTCTAACCTCTCAACTAATGAGATCACAATGATGTTAC	3506
Db	2367	GTACTCTACACTCTAAAGGAGTCATCATCATGTTGTTAGTGTGTT	2426	QY	3311	CGAACTAACTGGATGAGAACTCTAACCGAGTCACCGAGTCACACTGGGATA	3360
QY	2221	ATCGAAATGCCAACAGAGAGTCATCATCATGTTGTTAGTGTGTT	2280	Db	3507	CGAACTAACTGGATGAGAACTCTAACCGAGTCACCGAGTCACACTGGGATA	3566
Db	2427	ATCGAAATGCCAACAGAGAGTCATCATCATGTTGTTAGTGTGTT	2486	QY	3361	TCGGTGTAACTCTCICAACCTGGTACCGAGACTTCAGTAATGGAGTCTAGA	3420
QY	2281	TGAGAAGCTTCTGGATGACTCTCTGGTCATTAGGTTGAGGTTAAGGT	2340	Db	3557	TCGGTGTAACTCTCICAACCTGGTACCGAGACTTCAGTAATGGAGTCTAGA	3626
Db	2487	TGAGAAGCTTCTGGATGACTCTCTGGTCATTAGGTTAAGGT	2546	QY	3421	GCCTATACAAACACAGGACGCCCTCATCGGGAGGCCATAAGGCCT	3480
QY	2341	ACGTCCTGTTAGTGTGTTACCGGTAACTGAGGTTAATCTGAGATAAGT	2400	Db	3627	GCCTATACAAACACAGGACGCCCTCATCGGGAGGCCATAAGGCCT	3686
Db	2547	ACGTCCTGTTAGTGTGTTACCGGTAACTGAGGTTAATCTGAGATAAGT	2606	QY	3431	GGATCTCTGTCAGTTGGACTACACCAACCCGACTTCAGGAGTCACCTC	3540
QY	2401	GGTAAGGATCTTAACTGAGGTTAATCTGAGGTTAATCTGAGATAAGT	2460	Db	3687	GGATCTCTGTCAGTTGGACTACACCAACCCGACTTCAGGAGTCACCTC	3746
Db	2607	GGTAAGGATCTTAACTGAGGTTAATCTGAGGTTAATCTGAGATAAGT	2666	QY	3541	GGAGTCATAATGTCAGGACGATCTGGTGTGAGTTAAAGAACAGCTC	3600
QY	2461	GGCATACTCAGGAGTCATAATGCTCTGAGTCACACCTCTGTTAGTGTGAGT	2520	Db	3747	GGAGTCATAATGTCAGGACGATCTGGTGTGAGTTAAAGAACAGCTC	3806
Db	2667	GGCATACTCAGGAGTCATAATGCTCTGAGTCACACCTCTGTTAGTGTGAGT	2726	QY	3601	AGAGATGGCTACTCTGACTACCCAGAACTACTGTCAGGAGTCACCTC	3660
QY	2521	CGTATTGAGGAGACTATAGCGGCATATCTGCTGAGTTAATGGTCTCGCGAGGT	2580				

Db	3807	AGAAGATGGCACTCTGACTACCCAGAACTAGTGCTGAAGTGACAGCAATTC	3866	QY	4741	ATTCGACATAATAGGTTGGAATTGAGCACATCACTAGACCTCGAGTCGGC	4800
QY	3661	CTCGTAACTGCTAATAAGAATCGGGTAACTGAAAATAGATAGGGAAATATTC	3720	Db	4947	ATTCGACATAATAGGTTGGAATTGAGCACATCACTAGACCTCGAGTCGGC	5006
Db	3867	CTCGTAACTGCTAATAAGAATCGGGTAACTGAAAATAGATAGGGAAATATTC	3926	QY	4801	GTATCAAATGCTCTTATCTCATCCTAACCTAACTTATGCTTCCTCTAAATCAGT	4860
QY	3721	ATTCGCGGGGTGATGATGATTCTCGGGTAACTCAACTAAATGTCAGAACTAATG	3780	Db	5007	GTATCAAATGCTCTTATCTCATCCTAACCTAACTTATGCTTCCTCTAAATCAGT	5056
Db	3927	ATTCGCGGGGTGATGATGATTCTCGGGTAACTCAACTAAATGTCAGAACTAATG	3986	QY	4861	ACCTTGAGAAAAACATATGTTGTTATGAAAGGTTCTPATGGTTAAACAGTC	4920
QY	3781	TGAACAGGGGTGCTGAACTGAGTCAAGTCAGTCACTCAAGTTGAGGAGAAT	3840	Db	5067	ACCTTGAGAAAAACATATGTTGTTATGAAAGGTTCTPATGGTTAAACAGTC	5126
Db	3987	TGAACAGGGGTGCTGAACTGAGTCACTCAAGTTGAGGAGAAT	4046	QY	4921	GGTTTGGAAUTGTTACACATCAACATGAACTGAGTGGAGACATGGTTACTCAGT	4980
QY	3841	ACTTAATGACAACTGATATAGTCACTGTTGCTGTAAGTTGAGGAAAT	3900	Db	5127	GGTTTGGAAUTGTTACACATCAACATGAACTGAGTGGAGACATGGTTACTCAGT	5186
Db	4047	ACTTAATGACAACTGATATAGTCACTGTTGCTGTAAGTTGAGGAAAT	4105	QY	4981	TTACACATGAGTTTACATCTGATGAACTGAGTGGAGACATGGTTACTCAGT	5040
QY	3901	CTCTCCATTAAATCTTATGATGATGATCAGTCAGTCAGTCAGTCAGTCAGT	3960	Db	5187	TTACACATGAGTTTACATCTGATGAACTGAGTGGAGACATGGTTACTCAGT	5245
Db	4107	CCTCCATTAAATCTTATGATGATCAGTCAGTCAGTCAGTCAGTCAGTCAGT	4166	QY	5041	TTATATGTTGATGATCTATGTTGGAGTCAAAAGTTGATATTGT	5100
QY	3961	CAATTCACACAAATCGACATCAATCGAAATGAAATCACCAGTTAATGACA	4020	Db	5247	TTATATGTTGATGATCTATGTTGGAGTCAAAAGTTGATTAATTGT	5306
Db	4167	CAATTCACACAAATCGACATCAATGAAATGAAATCACCAGTTAATGACA	4226	QY	5101	GTATCAAATGAGACATTTGAAATTGAGATGTTGAAATATCAAATTCTTGG	5160
QY	4081	ATTGAAAGTATAATGTTGATGAAATTGAGATGTTGAAATATGAGATGACATGCA	4140	Db	5307	GTATCAAATGAGACATTTGAAATTGAGATGTTGAGACATGGTTACTCAGT	5356
Db	4287	ATTGAAAGTATAATGTTGATGAAATTGAGATGTTGAGACATGGTTACTCAGT	4346	QY	5221	ATTACTAAGGTTCAACTGAGTCACTCATATGGAAAGTTGAGTAAAGTTGATATTGT	5100
QY	4141	AGCTTAATCCACTACCGATGAAATATGTTGATGAAATGAGTTGAAATATAA	4200	Db	5427	ATTACTAAGGTTCAACTGAGTCACTCATATGGAAAGTTGAGTAAAGTTGATATTGT	5486
Db	4227	AAATACTGAAATATGTTGATGAAATGAGTTGAGATGTTGAGATGTTGAA	4286	QY	5281	TGCAAAATGAAAGGTTGCAATAATTCGTTGAAAGGTTATCGAGATGATTGCA	5340
QY	4347	ACCTAATCCACTACCGATGAAATGAGTTGAGATGTTGAGATGTTGAA	4406	Db	5487	TGCAAAATGAAAGGTTGCAATAATTCGTTGAAAGGTTATCGAGATGATTGCA	5456
QY	4201	TCATAGCATGTCAGGAAAGGAAATACATGTTGAGAAATGAGTTGAAAT	4260	QY	5341	AAAGGTTCCGATGAGACATGCTGCACTGTTGAAAGGTTATCGAGATGATTGCA	5400
Db	4407	TCATAGGATGTCAGGAAAGGAAATACATGTTGAGAAATGAGTTGAAAT	4466	Db	5547	AAAGGTTCCGATGAGACATGCTGCACTGTTGAAAGGTTATCGAGATGATTGCA	5656
QY	4261	TGCTGTTGAAACATCACTGTTGAGTCACTGTTGAGTCACTGTTGAGTTGAA	4320	QY	5401	TGCCCTCTTATGCGCACACACAGCTGTTGAAAGGTTATCGGTGTTAGTGA	5460
Db	4527	TGCTGTTGAAACATCACTGTTGAGTCACTGTTGAGTCACTGTTGAGTTGAA	4586	Db	5607	TGCCCTCTTATGCGCACACACAGCTGTTGAAAGGTTATCGGTGTTAGTGA	5656
QY	4381	AAACAAAGATGCAATCAAGTCACTGTTGAGTCACTGTTGAGTTGAGGTTAA	4440	QY	5461	GTGTTGTTCTAAATCGCAATCCAAATGTCAGTCACTGTTGAAAGGTTATCGGTGTT	5520
Db	4587	AAACAAAGATGCAATCAAGTCACTGTTGAGTCACTGTTGAGTTGAGGTTAA	4646	Db	5667	GTGTTGTTCTAAATCGCAATCCAAATGTCAGTCACTGTTGAAAGGTTATCGGTGTT	5726
QY	4441	CCCAAAGGGTAACTGCTGCAAAATGCACTATAATGATCCGGATGTT	4500	QY	5521	TAGTTATCACAAATGAGCTGGATATCTGTCAGTCACTGTTGAAAGGTTATCGGTGTT	5580
Db	4647	CCCAAAGGGTAACTGCTGCAAAATGCACTATAATGATCCGGATGTT	4706	Db	5727	TAGTTATCACAAATGAGCTGGATATCTGTCAGTCACTGTTGAAAGGTTATCGGTGTT	5786
QY	4501	AAAGTCACTGATGCTGAACTGAGAAATAGATCAAAAGTCTTAAAGAAGGTTCC	4560	QY	5581	ACCCCAAATCTTGTGTTATGCAATGTTGAGCTCATTCGACAGGATGCA	5640
Db	4707	AAAGTCACTGATGCTGAACTGAGAAATAGATCAAAAGTCTTAAAGAAGGTTCC	4766	Db	5787	ACCCCAAATCTTGTGTTATGCAATGTTGAGCTCATTCGACAGGATGCA	5846
QY	4561	ATTCCACCGCTGCAAGCTPATATCTCTGGTTGGACATAGTCAGTCAGTCAGTC	4620	QY	5641	TGAAATCTTGTGTTATGCAATGTTGAGCTCATTCGACAGGATGCA	5700
Db	4767	ATTCCACCGCTGCAAGCTPATATCTCTGGTTGGACATAGTCAGTCAGTCAGTC	4826	Db	5847	TGAAATCTTGTGTTATGCAATGTTGAGCTCATTCGACAGGATGCA	5906
QY	4621	TCTCAAGGTGTTGCTGGAATCAGTGTGTTGCTCCAGACAGCAAGGAA	4690	QY	5701	CAAAACAAAGGTTATGCAAGGCTCAGCAACTGTTGAGCTCATTCGACAGGATGCA	5760
Db	4827	TCTCAAGGTGTTGCTGGAATCAGTGTGTTGCTCCAGACAGCAAGGAA	4886	Db	5907	CAAAACAAAGGTTATGCAAGGCTCAGCAACTGTTGAGCTCATTCGACAGGATGCA	5966
QY	4681	ATTGGATTGACCTTTAGTGTAGTCAGCTGATGAGCTGACTATAAGAT	4740	QY	5761	TCAAGTGTGAACTGATGAAATAACATTAAATGATGTTGAGCTGAGTAG	5820
Db	4887	ATTGGATTGACCTTTAGTGTAGTCAGCTGACTATAAGAT	4946	Db	5967	TCAATGTTGAAAGCTATGCAATAAAACATTAAATGATGTTGAGTAG	6026

Db	4452	TATATTCATCA	4441	Db	301	ACCATTCGCCAAATAGCGTGATAATCAGTCAGTCAGTTATTGATT	360
RESULT	4			RESULT	5		
DEFINITION	AF030556	AF030555	408 bp	DEFINITION	DHA39551	DHA39551	DNA
ACCESSION	AF030555	Candida albicans	linear	ACCESSION	AU439551	3928 bp	linear
VERSION	AF030556.1	retrotransposon pCal LTR, and GAG gene, partial	PLN 01-JUL-1998	VERSION	AU439551.1	GI:3273502	PLN 07-JUL-2002
KEYWORDS		Candida albicans		KEYWORDS			
SOURCE		Candida albicans		ORGANISM			
ORGANISM		Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetales; Saccharomycetes;		ORGANISM			
REFERENCE		1 (bases 1 to 408)		REFERENCE			
AUTHORS		Goodwin,T.J.D. and Poulter,R.T.M.		AUTHORS			
TITLE		Temperature- and strain-dependent expression of the pCal		TITLE			
JOURNAL		retrotransposon of Candida albicans		JOURNAL			
REFERENCE		Unpublished		REFERENCE			
AUTHORS		2 (bases 1 to 408)		AUTHORS			
TITLE		Goodwin,T.J.D. and Poulter,R.T.M.		TITLE			
JOURNAL		Direct Submission		JOURNAL			
FEATURES	source	Submitted (21-OCT-1997) Department of Biochemistry, University of Otago, P.O. Box 500, Dunedin, New Zealand		FEATURES	source		
ORIGIN		Location/Qualifiers		ORIGIN			
Query Match		1. 408		Query Match			
Best Local Similarity		6.2%		Best Local Similarity			
Matches	403	Score 400; Pred. No. 1.1e-53;		Matches	403	Score 400; Pred. No. 1.1e-53;	
Conservative	0	DB 8: length 408;		Conservative	0	DB 8: length 408;	
Indels	5	DB 39: .398. ->408.		Indels	5	DB 39: .398. ->408.	
Gaps	0	/note="Start of GAG ORF"		Gaps	0	/note="Start of GAG ORF"	
DB		/codon_start=1		DB			
QY	1	/transl_table=12		QY	1	TGTGTTGACATTTGTTGAGAACTGATCATGAAATGATGTTATGAA	60
Db		/protein_id="GAG"		Db		TGTGTTGACATTTGTTGAGAACTGATCATGAAATGATGTTATGAA	60
QY	61	/protein_id="GAG"		QY	61	GATGGAATTTCATCACACATCAGGTAGAGAACTAACTATGTTAGT	120
Db		/db_xref="GI:3273503"		Db		GATGGAATTTCATCACACATCAGGTAGAGAACTAACTATGTTAGT	120
QY	121	/translation="MSSA"		QY	121	ATTAATAGGATGATAACATCCAGATAACAGATTCAGAGATAGAGGAGTT	180
Db				Db		ATTAATAGGATGATAACATCCAGATAACAGATTCAGAGATAGAGGAGTT	180
QY	121			QY	121	ATTAATAGGATGATAACATCCAGATAACAGATTCAGAGATAGAGGAGTT	180
Db				Db		ATTAATAGGATGATAACATCCAGATAACAGATTCAGAGATAGAGGAGTT	180
QY	181			QY	181	TCAATATATCTGTCTATTAATCTGCTCTATTCATAGAGTAAAGCGAT	240
Db				Db		TCAATATATCTGTCTATTAATCTGCTCTATTCATAGAGTAAAGCGAT	240
QY	241			QY	241	ACACGCTCAATCTGAGAAGTTATTCATAGAGTAAAGCGAT	300
Db				Db		ACACGCTCAATCTGAGAAGTTATTCATAGAGTAAAGCGAT	300
QY	301			QY	301	ATCATTCGICCCAAATTAGGTGATAATCGCTCTGAGATGTTATTGATT	360

MYHLMGNMISLESKMYLKSCHIKMSSRITASEBERVSCNECLAINSKOSSHINH
 FTAAPEFLFRHSDDTGIFHRKXYITLIDESGYANTINSEHKSQDILPEKIRI
 WNNKFEDVANIAFFTDNALEMPTDKOLABIGIRSESPRNGSERSRTSIIQF
 IRKALLP1ODTRTLYLPLKUVVYTRWTPRSKGCSCPMLYDNCKFYHNPQF
 GLD1-LVKLSSTVEAKRKYGSNSKSPC1FGTGTGTYGTYDHWVVKLTDKPI1VTP
 NITAMSKMDNKLVILKSLDYLQDDAODDISGKLTDRDKSIALEHMNDTIV
 CNEEDTOEVLNSUSTIOVAQHFERHESNSGMSPFVSHYTMVTDTSNDSPDYI
 BETTNEVNSNPDNNTEDSNKINFNDSNSNDSNQKSSTHOTLPHPAPDKD
 BINKLNEDGTHLSSKESSTSEKPTTQVRVTRSKKAVTVERGKTSKSR
 VKKPHKTSURNSSKSKINDLQSHAPKNSKETTYVTRSMRKNSKSK
 SLVTRKINTVTRKDITDNWQSIQREJTFKKEVTVKPKVNAKPTWVH
 KNDKKEVOKRSRVCWQGRQTAHEYDVKVSPVLSIRLMLTAVEWPIH
 LDTSAYAHADIDVPSKSTIVKPPGSNIDSCKWVNLNSVYGMKQAGTWWOCITKU
 MDLNEPDRDAISCHFCKYGENKLVIALYDMLFISNNTLNGELAKHDFK
 IFAD1-SEFQIEFVQIAGYRPLSHNFSVNLNSVCKLNUYNGY1PLKERNKSEI
 LKSEPTNDENVTKNISLQYDLSVNLNSVCKLNUYNGY1PLKERNKSEI
 SVDEBKLYLCKYKOTISFSLXKRNFSHKQSI1QTEDEGSAFPTKORR1CY
 TVIUANGNLINWSTRQKVITDSSACETKSRVLSRQALDNLVIDEITF
 EDNAVIANCNEETSYRSMWDFKLFKIRQVLYSEGILKLYVNTSINTADMILKALS
 RKLLENLSLIFENDNCE"

ORIGIN

LTR

Query Match 5.7%; Score 363.8; DB 8; Length 5928;
 Best Local Similarity 52.5%; Pred. No. 2.8e-48;
 Matches 872; Conservative 0; Mismatches 777; Indels 12; Gaps 3;
 QY 4444 CAGAGGTTTACATCGGCGTCACAAATCGACTATATGATCCGGATGATAA 4503
 Db 3824 CACTCGAAATAAATACTGTATATGAAAGTCAAGATAACATGGAAACCA 3983
 QY 4504 GTCATACTGAACTAGAGAACTTAGTCAAAGATGTTAGAAGAAGTCCATT 4563
 Db 3884 GTCATACTGAACTAGAGAACTTAGTCAAAGATGTTAGAAGAAGTCCATT 3943
 QY 4564 TCCACACGGTGTGAAACCTATACTATCGTGGGTACACTGAGAAATGTTCT 4423
 Db 3944 TCTTAAATACTGACAACTTACATCCAACTACTGTGGTCACTACATAAAATTAAAGATCT 4003
 QY 4624 CAAAGGPTGTTGTCGAAATCAGTGTGTGTCATGAGACAGACAGAACAGGAATT 4683
 Db 4004 CAAAGAAGTGTAGTAAATCAGTGTGGTGTAGGGCTTGGCAATTGCAATCA 4063
 QY 4684 GGATTATGACCTTGTAGTGTAGTGTACCTGTATAGTCTGTGACTATAGATT 4743
 Db 4064 ACACTATGATACTCGAGGTGTCACTCCGTGATGGATTCCATATCTTACT 4123
 QY 4744 GCAATATAGTGTGAAATTAGGATGACATTAAACATTAGCTGAGTGGCGTA 4803
 Db 4124 TACAGGATAGCAGTGAATGAAATGGCCGATCATCTTGATATCTCCGATA 4183
 QY 4804 TCTAAATGCTCTATACTCTCAATCAATTATGCTTCCTCTAACTGACC 4863
 Db 4184 TTACATGCACTATGCTGTTATGAAAGTGTCTATGCTAAACACCCCTGGATCGAA 4423
 QY 4864 TTGAGAAGAACCATGTTGTTATGAAAGTGTCTATGCTAAACAGTCGG 4923
 Db 4244 TATGATTCTGTAATGTTGCAATTAAACAAATCTGTGTTATGGAATGACAGCAG 4303
 QY 4924 TTGGATGTTGACACTATCAAAGAGATTTGGAGACATGGTTACTCAGTT 4983
 Db 4304 GATATGTTGATCAATGTTAATCTGTTATGGTCTTATTCGACCTGATAC 4363
 QY 4984 ACACATGATGTTTATCTGATGATGAAAGGATCAGTATATTTAGGTT 5043
 Db 4364 TGCATTTGAGGATCATTTGAACTTGTGAAATAGAGCTATGCTGACT 4423
 QY 5044 ATATGTGATGATCTTGTGAACTTGTGAAATTTGATATTTGTGAA 5103
 Db 5104 TCAATTGAGGATCATTTGAACTTGTGAAATTTGATATTTGTGAACTTGTGACT 4483
 QY 5104 TCAATTGAGGATCATTTGAACTTGTGAAATTTGATATTTGTGAACTTGTGACT 5163

Db 4484 TGAACCTGCTAAACATTGACCTTAATTCGCTGACATATCTGATTCTAGGAAT 4543
 QY 5164 TGATTGTTGTAACACCGATCTGTTATTATTCGAAAGAAATTCTCAAGAAATT 5223
 Db 4544 TGATTCATCAATTGCGAGGGGTATAGATPATCACACAACTTCTGATCTG 4603
 QY 5224 ACTTAAGGTTCAACTAGATGACTCTATGGAAACACATCCCTGGATCCGATG 5283
 Db 4604 AATTAATAATTCATTAACATAATTATGAAAGTATATCCATATTATTAAAGA 4660
 QY 5284 CAATATGAAAGTGTGAAATTCTGAAACAGTATTCGAGAGAAGTGTGAA 5343
 Db 4651 GAAATAATAATTTCGAA-----CTCACTAAGAGTGTGATTTACTATGACCAAGT 4714
 QY 5344 GGTCGCGATGAGACATGCTGACCCGTGCTAAAAACTATACAAAGTGTGTTGG 5403
 Db 4715 CATTAATAATGAACTATGCTTAATGAGAAGATAAAGACATATCATCAGGTTGG 4774
 QY 5404 CCTGCTTATGEGCTGCCAACACAGTCAGATATGAGTATACCGGTGATGTTGG 5463
 Db 4775 TTGTATATGGCAGCTATAACATGGCCAGACATAGTGTGAAATATT 4834
 QY 5454 GGGTTTAATCTGCAATCAATGTCATGATTAGGAATGATTTGGTTAG 5523
 Db 4835 GAGCTCAATAATCAAAATCCACATGTCATCTGAAATTATATTGTTACG 4894
 QY 5524 GATATCAATAGCATGGATATCAGTACATGAGTAAAGAACAGTCAATAC 5583
 Db 4895 TTGTAAACACACATATCTTTAGTTAGATAGTAAAGGAAATGCAATTAA 4954
 QY 5584 ACCAAATCATTTGTTGAAATGTTGTCAGTGTGATGCTCATTGCAACAGGATGGATG 5643
 Db 4955 AAGGGTCAATTATCCTCAACTTTCTGATGTTGCTTGTGACCCACAAAGAT 5014
 QY 5644 AAAATCTTGTGAACTTGTGATGGAAATTGGTCAATGGCGACAA 5703
 Db 5015 AAGCTGTTACAGGTATACCTGATGTTTAAATGGTATCTAATAAATTGCTTACCAT 5074
 QY 5704 AAACAAACGGCCTACACAAAGCTCAGCAGCTGTCAGCAGTGTGCTTAATATAC 5763
 Db 5075 AAGCAGAGGTATACCTGATGTTGCTGAGCTTGTGAAATCACTGACATCAGC 5134
 QY 5764 AATGTGAAAGTATGAAATAAAACACATTAAATGATGGTTGTGAGTAGTAA 5823
 Db 5115 TGTGAGAGTACCTAAATCAGACAGTATATGATTAATTTAGTATGAGTAA 5194
 QY 5824 GATACATTTGTCATCAAGAACCAAGTGTGTTAGTTGAGAAATA--CTATG 5880
 Db 5195 AATACATATTGAAAGAACGCCGAGTATAGGAACTGTAAATAGBAGGFACTC 5254
 QY 5881 TACCCCACTCACCATAAGATACTGCTATAGTTCTAGCCATTGATCAATGATAA 5940
 Db 5255 ATATCAAGGCTATGGTGTGATCAATTAAATTTACAGACAGTTCTGAGG 5314
 QY 5941 AATTTTCAATCTCTATGAGCAAGATGATATTGGCGGATGTTGACTAGTG 6000
 Db 5315 TATTTAAATTAATGTTGAACTGATGTTGAGATGTTACAGACAGTCACAGGC 5374
 QY 6001 TCAACTGCTGCTAAATCAAGCATCTGTTGAGGCTATGATAAACGGTAGCTAGA 6060
 Db 5375 ATTAAGCCGAAATTATTTGAAATCTCTGTTCACTACTATTGAAAGAATGATATAA 5434
 QY 6011 AGATAATCAACTGATCAAAATGCAATACGGCAGAT 6101
 Do 5435 TAAGAATAGAACTTACTATGTCAGAGTACCAAT 5475
 RESULT 6
 A078809 AF078809
 DEFINITION *Candida albicans* Tca4 retrotransposon reverse transcriptase (pol)
 ACCESSION AF078809

VERSION AF078809.1 GI:3377678
 KEYWORDS
 SOURCE Candida albicans
 ORGANISM Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; mitosporic
 REFERENCE 1 (bases 1 to 1470)
 AUTHORS Goodwin, T.J.D.
 TITLE Tca4, a Ty1-copia retrotransposon from *Candida albicans*
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1470)
 AUTHORS Goodwin, T.J.D.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-1998) Department of Biochemistry, University of Otago, Cumberland Street, Dunedin, New Zealand
 FEATURES source
 /location/Qualifiers
 1. 1470
 /organism="Candida albicans"
 /mol_type="genomic DNA"
 /strain="SC5314"
 /db_xref="taxon:5476"
 <1..>1470
 /gene=pol
 <1..>1470
 /gene="pol"
 /codon_start=2
 /transl_table=12
 /product="reverse transcriptase"
 /protein_id="AAC2090.1"
 /db_xref="GI:3377679"
 /translation="GRSVNPPKPVLAIVKKDYNQKEWROSNEEIEKFKANQVTV
 EKTRPNVWPLKTMWHTYKNDIKNHNYKSRCTVNGNUNVERDPAHAISSPVW
 SIRLSAIAVENNUWQDLSKASVYNSDCEVIFRPPRGREVERGYSRHLVST
 YGLQSARAHWYSHKRNVBLANGKQTLNDGILWENYDLYVSYVDDPFLMHN
 EKTKFKEVMTETYFOLOFGFATEYLGQFQPKDGTDLDFPLKVLVAFNFIODS
 YGKDFLIPKIDINWQDLSKQNDFVKQKLINSVSKYIDYENHEHP
 ESKQPEPLSKRGKILYQSAVGSLLWASMTRPDLAFSINQLGKCARPDVDPWKRUM
 CLRYKDDMDPKLUVKRGILNGKSDFTBCEFDASAPDURKSINGTSIVNGNLV
 MATKQKIIHSSARACMAMLYNTIKAFDURNTEDL"
 repeat_region
 /note="Tca4 retrotransposon"
 /rpt_family="Ty-1-copia"
 /rpt_type=dispersed

ORIGIN

Query Match 5.3%; Score 3392; DB 8; Length 1470;
 Best Local Similarity 54.4%; Pred. No. 3.4e-44; Mismatches 790; Conservative 0; Indels 84; Gaps 2;
 Matches 790; Conservation 0; Mismatches 78; Indels 84; Gaps 2;

QY 4423 TCTAGAGCGGTATAACCCAGAGGTTACATGCGTCACAAATGACTA 4482
 Db 1 TGGAGATCAGTCATCCACCTAAAGATATTAAATGCAATTGTTAAAGAAATGATTA 60
 QY 4483 TAATGATCGGGATGGTAAACTCAATGATGCTGACTAGAGAAATTGATCAAGA 4542.
 Db 61 CAACTAAAGAATGGCTCAAGATGAGAAGAAATGAAATTAGGCTAACCA 120
 QY 4543 TGTTCACAGAGATGCCAATGCCACGGGTGGAACGCTATATCTGGTTGGTCA 4602
 Db 121 AGTTTACCGTGTAAACACCAAAAGAACGCTTGTCCATTGAAACCTGTTGAGTCA 180
 QY 4603 TACTGAGAAATGATCTCTCAAGGTTGTTGCGAAATCAGGTGTGCGATG 4662
 Db 181 TACATACAAACCAATGAGCTCAAACATATAACAAAGCGTGTGCGGTATGG 240
 QY 4663 CAACAGACAAAGGAAATGGATGACCTTGTGTTAGTCACTGTGTTAGA 4722
 Db 241 AACATATATGGTGTGAAATCGGATTTGATGCCATGCCATCTCTCCCGGTAGTAA 300
 QY 4723 TCTGTGACTAGATAGATAGATAGTTGATAGTAGGATGACATCAACAA 4782
 Db 301 TCTCACAGTATAGCAGCTTCTGCCATAGCTGTGAAATAACTTGGTTATGCACCA 360
 QY 5779 CGAAATAAAAA 5790

QY 4783 TTAGACGTCAGTCGGCATCTAAATGCTCTTACTCAATCAACCAATTATGT 4842
 Db 361 ATGGACATGCCCTCGCTTATTCGACGCCAGTTGGAGATGGAGAAATCTTGT 4902
 QY 4843 CTTCTCTTAATCTGTTACCTTGAGAAACCATTTGTTGAGATGAAACCTTGT 4902
 Db 421 GAGACCCACCGCTGTTTGAGTTAACCTGGCTTATGTTGCGTTACAAAGCTGT 480
 QY 4903 CTATGGTTAAACACTCGGTTGAAATGGTATCACATCAACAGATGGAGA 4902
 Db 481 GTACGCTCTAGCAGGTGCCATATGGTAACTGGTAACTGGTAACTGGAGA 500
 QY 4963 CATGGTTTACTCAAGTTACACATGATGTTGTTGAAAGAGGG 5022
 Db 541 AAATGTTAAACACACTACATGATGTTGCTTGTGAAATATGAAATG 500
 QY 5023 ATCAGTATATATTAGGTTATGTTGATGATCTTATGGTTGGAGTACCAAA 5082
 Db 601 AGATGATATATGTTGAGTGTATAATGTTGATGTTTATGATGATGAA 660
 QY 5083 ACTTATGATATTGTTGGTATGATGAGATCATTTGAGTAAAGTTGTTGCA 5142
 Db 661 GATTATTAAGGTTGTTGCTATGCGAACATTTTCAGTACATATTGCGA 720
 QY 5143 AATATCAATATCTGGTATGAAATTCGTAACACGAATCTGTTATTTCTCA 5202
 Db 721 GCTCTGGAGATCTGGAAATGAGTAACTTGTGAAAGCCTGATACAGTGTGACCA 780
 QY 5203 AGAAAATTCTCAAGAAATTACTAAGGTTCAACTAGATGCTCATGGCAA 5262
 Db 781 ATTACCATCTCTGGAAATATGGTCTTACATTCATATTCAGATAGTTGAAAGA 840
 QY 5263 CATACTCTGGATCTCCAAATGCAAAATG 5291
 Db 841 CATCTCTTAAATCCCAAAGATATTATGTCAGTCAGCAATGAGAACTGCAATG 900
 QY 5292 AAAGGTTGCAATAATCTGTCAAAC 5317
 Db 901 AATGATTTCTGTTAAGTGGAAACTCTAAAGTTATGAACTGGGTCAAACCA 960
 QY 5318 -----GTTATCAGAGATGTTGAAANGTTCGAATGAGAC 5358
 Db 961 GTCATGATGATTACGAAAGAAATGAAAGACATGTTAAAGATCACCAAGCTGA 1020
 QY 5359 ATGGCTTGACCTGTGACTAAACATATACCAACGGTTGCCTGCTTATGGC 5418
 Db 1021 ACCTTAAGCAGAAAGAAATATACCAATCGCAGTGGGACTGTTGTTATGGC 1080
 QY 5419 TGCACAAACACACGTCAGATATCGGTTGAGTCTGTTGGTCTAACTGC 5478
 Db 1081 TTGCTGAACTAGGCCAGACTTACCTTGTGAACTCAGCTGGACAAATGTC 1140
 QY 5479 AAATCCAAATGTCATGTTATGAGAAATTGTTGTTAGGATATCAAAATAG 5538
 Db 1141 TCAACAGATGTCAGACTGGAGAGATGTTGTTGCTACGGTATTTAGAAGA 1200
 QY 5539 CATGGGATATCAGTGTGAAATGAGATGAAATACCAAAATCATCTGT 5598
 Db 1201 TATGGACTCAACTATGATACAGAGAGACTCACACAAAGTCACAGATTAT 1260
 QY 5599 TATGGATGTTGAGTGTGCTCATTCAGGATGAAATCTATGTTGTT 5658
 Db 1261 TATGGATGTTCTCAAGTGTGCTCATTCGGCCAGATGGACAGAACTCAACCGG 1320
 QY 5659 AACTTGTTATGAGATGAAATTGGTGCATGGGACAAACAAACAGCGTCAT 5718
 Db 1321 GACAGCTTGTGAAATGGCAATCTGTCATGGCAACAGAACAAATAAT 1380
 QY 5719 AGCAACAGCTCGAGCTGTAATGTTGCTCAATTATACATGTTGAACCTAT 5778
 Db 1381 CACAGTATCTGAGTGTGCAATTAATTACATGTTGAACCTAT 1440
 QY 5779 CGAAATAAAAA 5790

REFERENCE	1	4830	ATCCAATTATGCTTCTCTTAATCAGTACCTTGAGAAGAACCATGTTGTT	4830
AUTHORS	Shauil,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,	TGTTTGTGTCACACCCAGATTGAAATGCTGGAGAAGAAATATGTTAAC	69753	
TITLE	Cook,D., Kim,D. and Roe,B.A.	4890		
JOURNAL	Medicago truncatula BAC Clone mth2-54a24	TGAAACGTTCTGCTATGGGTTAACAGTGGTTGAAATGCTACACTACAA	4919	
UNPUBLISHED	Unpublished	4919	4919	4919
REFERENCE	2	(bases 1 to 128101)	(bases 1 to 128101)	4919
AUTHORS	Shauil,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,	TCACACAAACACTATGGTTAACAGTGGTTGAAATGCTACACTACAA	69753	
TITLE	Cook,D., Kim,D. and Roe,B.A.	4950		
JOURNAL	Direct Submission	GAGTATGGAGACATGTTTACAGTTTACACANTGATGCTTATTACATTG	5009	
COMMENT	Submitted (22-FEB-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	5009	5009	5009
-----	On Dec 12, 2003 this sequence version replaced gi:38678580.	5010	5010	5010
Center: Department Of Chemistry And Biochemistry	The University Of Oklahoma	AATATGAGAGGGATGAGTATATGTTATGTTATATGTTGAGATATCTTGTG	5010	5010
Center code:UNXOR	Center code:UNXOR	AGCTTACATGAGGAGCTGTTATGTTATGTTGAGATATCTTGTG	5010	5010
-----	-----	5010	5010	5010
* NOTE: This is a 'working draft' sequence. It currently	* NOTE: This is a 'working draft' sequence. It currently	5010	5010	5010
* consists of 2 contigs. The true order of the pieces	* consists of 2 contigs. The true order of the pieces	5010	5010	5010
* is not known and their order in this sequence record is	* is not known and their order in this sequence record is	5010	5010	5010
* arbitrary. Gaps between the contigs are represented as	* arbitrary. Gaps between the contigs are represented as	5010	5010	5010
* runs of N, but the exact sizes of the gaps are unknown.	* runs of N, but the exact sizes of the gaps are unknown.	5010	5010	5010
* This record will be updated with the finished sequence.	* This record will be updated with the finished sequence.	5010	5010	5010
* as soon as it is available and the accession number will	* as soon as it is available and the accession number will	5010	5010	5010
* be preserved.	* be preserved.	5010	5010	5010
* 1	14715: contig of 14715 bp in length	5010	5010	5010
* 14715	14815: gap of unknown length	5010	5010	5010
* 14815	128101: contig of 113286 bp in length.	5010	5010	5010
FEATURES	Location Qualifiers	ATTTTATCAGAAATTC	5213	5213
source	1. 128101	70113	TAGTGAGTATGAGAAATAC	70136
-----	-----	70113	TAGTGAGTATGAGAAATAC	70136
ORIGIN	Medicago truncatula BAC library H2"	70113	TAGTGAGTATGAGAAATAC	70136
Query Match	1.6%; Score 105.6; DB 2; Length 128101;	70113	TAGTGAGTATGAGAAATAC	70136
Best Local Similarity	47.3%; O: Mismatches 352; Conservative	70113	TAGTGAGTATGAGAAATAC	70136
Matches	389; Indels 3; Gaps 1;	70113	TAGTGAGTATGAGAAATAC	70136
QY	4470 ACAAATGACTATAATGATCCGGATGGATAAGTCATGATGTTGAACTAGAGAAT	70113	TAGTGAGTATGAGAAATAC	70136
Db	69396 ACAAAGAACGATGAGACTATGTTGAGAACTAGTGTGAGACTACATCAA	70113	TAGTGAGTATGAGAAATAC	70136
QY	4530 TTGATCAAAAGTTGTTACGAGAGTTCCATTCCACCGGTGAGACCTATCTA	70113	TAGTGAGTATGAGAAATAC	70136
Db	69456 TTGAGAGAAATCAACCTGAGAGCTGTTACTGCGAGACAGAAAGAGATGATG	70113	TAGTGAGTATGAGAAATAC	70136
QY	4590 TGGTGGTGTACATCTGAGAAATGATTCTCAAGGGTTGTCGAAATCGTT	70113	TAGTGAGTATGAGAAATAC	70136
Db	69516 TCAATGGGTTTCTGAGTGAATTGACTGATGCTAACTGAAAGCAA	70113	TAGTGAGTATGAGAAATAC	70136
QY	4650 GTGTTGTCATGGCACAGACAGAAATTGATTATGCTTACCCATTGTTAGT	70113	TAGTGAGTATGAGAAATAC	70136
Db	69576 GATGTTGCTCTAGAGTTCTCTGAGACAA--GACTTAGACTATAAGCTT	70113	TAGTGAGTATGAGAAATAC	70136
QY	4710 CACCTGTTATGATCTCTGACTATAAGTATTGACATAATAGCTTGAAATAGAA	70113	TAGTGAGTATGAGAAATAC	70136
Db	69633 CTCGAGTCTAGATGAGAACCTGAGTTGAGATGAAT	70113	TAGTGAGTATGAGAAATAC	70136
QY	4770 TGACATTCACAACTTACAGCTGAGTCGGCTATCTAAAGCTCTATTACTCTCA	70113	TAGTGAGTATGAGAAATAC	70136
Db	69693 GGTCACTTACCAATTAGTGTAGTCAGTCACTTAAATGTCCTAGAGGGTAG	70113	TAGTGAGTATGAGAAATAC	70136
QY	4795 * NOTE: This is a 'working draft' sequence. It currently	70113	TAGTGAGTATGAGAAATAC	70136
Db	* consists of 2 contigs. The true order of the pieces	70113	TAGTGAGTATGAGAAATAC	70136
QY	* is not known and their order in this sequence record is	70113	TAGTGAGTATGAGAAATAC	70136
Db	* arbitrary. Gaps between the contigs are represented as	70113	TAGTGAGTATGAGAAATAC	70136
QY	* runs of N, but the exact sizes of the gaps are unknown.	70113	TAGTGAGTATGAGAAATAC	70136

AP006375/C	AP006375	LOCUS	7280 bp	DNA	linear	PLN 22-JUL-2003
		DEFINITION	Lotus corniculatus var. japonicus genomic DNA, Chromosome 4, clone:LT32M04, TM0219, complete sequence.			
		ACCESSION	AP006375			
		VERSION	AP006375.1			
		KEYWORDS	HTG.			
		SOURCE	Lotus corniculatus var. japonicus (Lotus japonicus)			
		ORGANISM	Lotus corniculatus var. japonicus			
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.				
		REFERENCE	Asamizu,E., Kato,T., Sato,S., Nakamura,Y., Kaneko,T. and Tabata,S.			
		AUTHORS				
		TITLE	Structural Analysis of a Lotus japonicus Genome. IV. Sequence Features and Mapping of Seventy-Three TAC Clones which Cover the 7.5 Mb Regions of the Genome			
		JOURNAL	Department of Plant Gene Research; 2-6-7 Kusuma-kamatachi, Kisarazu, Chiba 292-0810, Japan (E-mail:ssatcc@kazusa.or.jp)			
		REVISION	URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex. 2337), Fax: 81-438-52-3934)			
		AUTHORS	Sato,S.			
		TITLE	Direct Submission			
		JOURNAL	Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute, Department of Plant Gene Research; 2-6-7 Kusuma-kamatachi, Kisarazu, Chiba 292-0810, Japan (E-mail:ssatcc@kazusa.or.jp)			
		REVISION	URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex. 2337), Fax: 81-438-52-3934)			
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				
		REVISION				
		AUTHORS				
		TITLE				
		JOURNAL				

Db	56113	GAAAAGGCTCTGTATGGCTAAACAGACCACTAGAGGATGGATGAAAGATTAACATA	56054	Query Match 1 6%; Score 105 4; DB 8; Length 127731;
Qy	4951	AGTATGGAGAGACATGGTTTACCGAACGTTTACAGATGATGGTTTATTCACATGAA	5010	Best Local Similarity 47.0%; Pred. No. 1.3e-07; Mismatches 441; Indels 6; Gaps 2; Matches 396; Conservative 0; MisMatches 441; Indels 6; Gaps 2;
Db	56053	ATTCCTGACCAACAAATGGTATGACAAAGCTGGCATGAAAGACCTGTTGTTAAGAA	55994	Qy 4471 CAAATGACTATAATGATCCGGATGGAAGTCATGAACTGACTAGAGAAATT 4530
Qy	5011	ATATGAGAGGATCTGTTATATTAGGTTATGTTGATATCTTATGGTGG	5070	Db 30057 CAAAGAGCTGTGACTGTGATGAGTGTCTGGATTCATCCATGAAAGGAAGCTGGACAGT 29998
Db	55993	AAATGCGACTTATGTTAGTCAA--ATTATGTTGATGACATGTTGGG	55937	Qy 4531 TAGATAAAGAGTTACGAGAACTTGGACTTGGAAACAGATGAAATCTGTTAAGGTA 5130
Qy	5071	AGTTGACAAAGTATTGATAATTGTTGGATCATGAGAGATCATTGGAGTAA	5130	Db 29997 CAAAGAGATGAGTGGGATGGTCCAAAGACCTGTGATGCAATGTTGGGAC 29938
Db	55936	CATGCGAACAGATGTTGACAAATTGGTGAACAAAGTGAATCTGAGTTGAATGAA	55877	Qy 4591 GGGTGGGTACATCTGGAGAAATTGTTGCTCAAAAGGGTGTGTTGGAAATCAGTG 4650
Qy	5131	AGTGTGTTGCAATTCATAATTCTTGTGTTGAGGATTTGTTGAGAATCTGAGTTGAA	5190	Db 29937 TAATGATATCAGGACACAGTCGATGAGTGAAGGTTAATGAACTGTTGAGATC 29878
Db	55876	TCTGTTGTTGAAATTCTGTTGAGGAAATTCTTGGAGATTCTGAGTTGAGTAA	55817	Qy 4651 TGTGTCATGCAACAGACAAAGAAATTGGGATATGACCTTGTGTTAGTTGTTGTC 4710
Qy	5191	TATTTATCTCAAGAAATTCCTGAGAAATTCTTGGAGATTCTCAACTAGTGASTC	5250	Db 29877 TCTAGT--TGCTCAGGGATCTACTCTCAATAGAAGGAATGACTTATGAACTTTC 29821
Db	55816	ATTCACTCACAGAAGTAGTGTGTTGAGGATTTGTTGAGTTGAGATC	55757	Qy 4711 ACTGTGATAGATCTGTGATGATAGATAGATGATGATGATGATGATGATGATG 4770
Qy	5251	ATATGGAAACATACCTGATTCGATGACAATATGAAAGGTTGCAATATTG	5310	Db 29820 TCTCTGTTGCTGCTGTTGATGATCTGGTGTGTTGGGATGTCATGCTTAA 29761
Db	55756	AGGTCAACAAAGAACCTGCTGCAACTCATATCACGTCACAAAGGATGAGAAGGTAC	55697	Qy 4771 GACAATCACTTGAAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 4830
Qy	5311	TGA 5313	Db 29760 CAGATATATGATGATGATGTCAGATGTCATTCATCTGAATGTTACTTGATGAGAAGT 29701	
Db	55696	TGA 55694	Qy 4831 TCCAAATTATGTCCTCTCTAAATCAGTACCTTGGATGGATGGTACACTATCAAAAG 4950	
RESULT 12			Db 29760 CAGATATATGATGATGATGTCAGATGTCATTCATCTGAATGTTACTTGATGAGAAGT 29701	
AP00094/c			Qy 4891 GAACGCTCTGCTATGGTAAAGAGTGGTTGAGTCAAGTACGTTGGATGGTACACTATCAAAAG 4950	
LOCUS	AP006094	127731 bp DNA linear PLN 22-JUL-2003	Db 29640 GAAAGGCTCTGTATGGCTAAACAGACCACTTGGAGCTTGGATGAAGATTAACAGA 29581	
DEFINITION	Lotus corniculatus var. Japonicus genomic DNA, chromosome 4,		Qy 4951 AGTATGGAGAGACATGGTTTACTCTGGTTTACAACTGATGATGATGATGATGATG 5010	
ACCESSION	Clone:J1T39H01, TM0172, complete sequence.		Db 29580 ATTCCTGATCACAGATGGTATGACAAAGGGCTGTGACAGACCTGTTGTTAGAA 29521	
VERSION	AP006094		Qy 5011 ATATGAGAGGATCTGTTATATTGTTGATGATGATGATGATGATGATGATGATG 5070	
KEYWORDS	HTG-		Db 29520 AAATGGCGGTGACTTATGGTACGTCAA--ATTATGTTGATGATGATGTTGTTGG 29464	
SOURCE	Lotus corniculatus var. japonicus (lotus japonicus)		Qy 5071 AGTTGACAAAGTATTGATAATTGGATGATGATGATGATGATGATGATGATGATG 5130	
ORGANISM	Lotus corniculatus var. Japonicus		Db 29463 CATGCTGAACTGGACAACTTGGTGAACAACTTGGTGAACAACTTGGTGAACAA 29404	
HTG-	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		Qy 5131 AGTGTGTTGTTGAAATTCCTAAATTATGTTGATGATGATGATGATGATGATG 5190	
KEYWORDS	Spermatophyta; Magnoliophyta; eudicots; core eudicots; roids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.		Db 29403 TCTGTTGAGGTATTACTTCTGGCTTGGATGATGATGATGATGATGATGATG 29344	
REFERENCE			Qy 5191 TATTTATCTCAAGAAATTCCTGAGAAATTCTGAGAAATTCTGAGAAATTCTGAGTC 5250	
AUTHORS	Kaneko,T., Asamizu,E., Kato,T., Sato,S., Nakamura,Y. and Tabata,S.		Db 29343 ATTCACTCACACAGAAGTAGTGTGCAACTTGGGATGTTGAGATGTC 29284	
TITLE	Structural analysis of a Lotus Japonicus genome. II. Sequence features and mapping of sixty-two TAC clones which cover the 6.7 Mb regions of the genome		Qy 5251 ATATGGAAACATACCTGATGACAATATGAAAGGTTGCAATTATG 5310	
JOURNAL	DNA Res. 10 (1), 27-33 (2003)		Db 29283 TGGTCACAAAGAACCTGTCGCAACTCATACGTCACAGAGATGAGATC 29224	
MEDLINE			Qy 5311 TGA 5313	
PUBMED	12693552		Db 29223 TGA 29221	
REFERENCE	2 (bases 1 to 127731)			
AUTHORS	Sato,S.			
TITLE	Direct submission			
JOURNAL	Submitted (12-DEC-2002) Shusei Sato, Kazusa DNA Research Institute, Department of Plant Gene Research, 2-6-7 Kazusa-kamatai, Kisarazu, Chiba 29283, Japan (E-mail: sato@kazusa.or.jp', Tel: 81-438-52-3935(ex. 2337), URL: http://www.kazusa.or.jp/, Fax: 81-438-52-3934),			
FEATURES	Location/Qualifiers			
source	1. 127731 /organism="Lotus corniculatus var. japonicus" /mol-type="genomic DNA" /variety="japonicus" /db_xref="taxon:34305" /chromosome="4" /clone="J1T39H01" /clone_id="J1T39H01" /note="TAC clone:TM0172-synonym: Lotus japonicus"			
RESULT 13				
AP006381	AP006381	105867 bp DNA linear PLN 22-JUL-2003		
LOCUS	Lotus corniculatus var. Japonicus genomic DNA, chromosome 4,			
DEFINITION	Clone:J1T09G09, TM0234, complete sequence.			
ACCESSION	AP006381			
VERSION	AP006381.1			
KEYWORDS	HTG-			
ORIGIN				

SOURCE	Lotus corniculatus var. japonicus (<i>Lotus japonicus</i>)
ORGANISM	Bularyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; roids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lous.
REFERENCE	1 Asamizu,E., Kato,T., Sato,S., Nakamura,Y., Kaneko,T. and Tabata,S. Structural Analysis of a <i>Lotus japonicus</i> Genome. IV. Sequence Features and Mapping of seventy-three TAC clones which cover the 7.5 Mb Regions of the Genome. DNA Res. (2003) In press
JOURNAL	2 (bases 1 to 105887)
AUTHORS	Sato,S.
TITLE	Direct Submission
JOURNAL	Submitted (01-MAY-2003) Shusei Sato, Kazusa DNA Research Institute, Department of Plant Gene Research, 2-6-7 Kazusa-kamattari, Kisarazu, Chiba 292-0818, Japan (E-mail:sato@kazusa.or.jp, Tel:81-438-52-3935(ex. 2337), Fax:81-438-52-3934)
FEATURES	Location/Qualifiers
source	<p>1. 105887 /organism="Lotus corniculatus var. japonicus" /mol_type="genomic DNA" /db_xref="taxon=34305" /chromosome="" /clone="J1b="IJT library" /clone="J1b="IJT library" note="TAC clone:TM0234-synonym: Lotus japonicus"</p>
ORIGIN	Query Match 1.6%; Score 103.8; DB 8; Length 105887; Matches 395; Conservative 46.9%; Pred. No. 2.5e-07; 0; Mismatches 442; Indels 6; Gaps 2;
QY	4471 CAATTGCACTATAATGATCCGGATGATAAAGTCATGATCTGAACTAGAAATT 4530
Db	89802 CAAGAGAGCTGACTGATGAGTCTGGATTCAATCCATCGAAGGAGACTGGACAGTT 69861
QY	4531 TAGATCAAAGATGTTAGGAGAGTCCATTCCACCGGTGAGCCATTATCT 4590
Db	89862 CAAGAGAAATGATGTTGGATTGGTCCAGACCTGTGATGCAATGTTGGAC 89921
QY	4591 CGGTGGTAGTACTAGAGAAATGATCTCTCAAGGTTGTCGAAATCACGTG 4650
Db	89922 TAAGTGATATTCAGAACAGTCATGAGAGTGTATGACAGAACAAATCTG 89981
QY	4651 TGTGTGCCATGCCAACAGACAAAGAAATTGGATTGACCTTGTAGTTAGTC 4710
Db	89982 TCTAGT---TCTCAGGGAACTCTCAATAGAGAGGGTAGACTTGTGAAACATTGC 90038
QY	4711 ACCCTGTTATGATCTGGACTATAGATTAATGCAATATAGTTGGAAATTAGGT 4770
Db	90039 TCTCTGTCGCTGCTGATCTGATGTTGAGGATGATGTTGCTAAATT 90098
QY	4771 GACAATTCACATTAGACGGAGCTCGGGTGTGTTACTATTCAA 4830
Db	90099 CAGACTATATAATGAGCTGAGGTGATCTGATGTTACTATGAGAACT 90158
QY	4831 TCCAAATTATGCTTCTCTTAATCAGTCCTTGAAGAAACCATGTTGTTATT 4890
Db	90159 CTATGTTGAAACACCTAAAGGGTTAGATCCAAAGTTTCAGATCTGTGACATT 90218
QY	4891 GAACTTCTGCTCTAGGGTTAACAGCTGGTTGGAGGTGAGCTACATCAAG 4950
Db	90219 GAAAGAGGATCTGTTAGCTAAACAGCTAGAGCTGTTAGAAGATGACAGA 90278
QY	4951 AGTATGGAAACATGGTTTACTCAAGTTAACACATGGTTTACATGAA 5010
Db	90279 GTTCTGATGACAACTGGTTAGCAAGCTGGCATTCAGAACCTTCTTCAAG 90338
QY	5011 ATATGAGAGGATCAGTAATATATTAGGTTATATGATATCTTATGGTGG 5070
RESULT	14
AC	AC137081
LOCUS	AC137081 129511 bp DNA
DEFINITION	LINEAR WORKING DRAFT
VERSION	HTG 12-SEP-2003
KEYWORDS	AC137081_1.9 GI:34581673
REFERENCE	HTG; HGS_PHASE1; HGS_DRAFT
SOURCE	Medicago truncatula (barrel medic)
ORGANISM	Bularyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; roids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliace; Medicago.
REFERENCE	1 (bases 1 to 129511)
AUTHORS	Cook,D., Kim,D. and Roe,B.A.
TITLE	Medicago truncatula BAC Clone
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 129511)
AUTHORS	Shailly,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE	3 (bases 1 to 129511)
AUTHORS	Shailly,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (12-SEP-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
COMMENT	On Sep 12, 2003 this sequence version replaced gi:3490629.
Center: Department Of Chemistry And Biochemistry	Center: Department Of Chemistry And Biochemistry
Center code:OKNOR	Center code:OKNOR
<p>* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.</p> <p>* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.</p> <p>* be preserved</p> <p>* 1 1159: contig of 1159 bp in length</p> <p>* 1 1169: gap of unknown length</p>	

FEATURES		RESULT	15
source		AC135396/C	AC135395
ORIGIN		LOCUS	Medicago truncatula clone mth2-33018
Query Match	1.6%; Score 103.4; DB 2; Length 129511;	DEFINITION	unordered pieces.
Best Local Similarity	46.8%; Pred. No. 2.7e-07; Mismatches 436; Indels 6; Gaps 2;	ACCESSION	AC135396
Matches 391; Conservative 0; Mismatches		VERSION	AC135396.25 GI:39841085
Qy	4471 CAAATGCACTATAATGATCCGGATGATAAGTCAAGAATCTGAATAGAAATT 4530	SOURCE	HTG: HTGS PHASE1: HTGS DRAFT,
Db	50154 CAAAGAGCCCTAACGTGATGTTGAGCTTGTGAACTGAACTAAGATCAAT 50213	ORGANISM	Medicago truncatula (barrel medic)
Qy	4531 TAGATCAAAGATGTTACGAGAAGTCCATCCACCGGTGAACTATCTAT 4590	KEYWORDS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; euroids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
Db	50214 CAAAGAGGTCAGTGTGAGATTAGTTCAAGACCGGAGGTATAAATGATAGGTAC 50273	REFERENCE	1 (bases 1 to 134544)
Qy	4591 GGGTGGGACTACATGAGAAATGATTCTCTCAAGCTGTTGTCGGAAATGAGTTG 4650	AUTHORS	Shailly,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
Db	50274 CAACTGGACTCTAAAGAAATCTGAGA--AATGACTGTGACAGAACAGGC 50330	TITLE	Medicago truncatula BAC Clone mth2-33018
Qy	4651 TGTGTTGCACTGCAAGACAAAGAAATGGAATTGACCTTGTAGTTGTTAGTC 4710	JOURNAL	Unpublished
Db	50331 AAGATTAATAGTTAAAGCTACACACATATGAGGGCTGAGCTGACTTGTGAC 50390	REFERENCE	2 (bases 1 to 134544)
Qy	4711 ACCTGGTATAGCTGTGACTATAAGATAATGACAAATAGTTGAAATTAGGAA 4770	AUTHORS	Shailly,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
Db	50391 TCTCTGTTGCTAGTTGATATAAGAATATTGCTAGTGTATCATGTTACTCAAGT 50450	TITLE	Direct Submission
Qy	4771 GCAATTCACATTGAGCTGAGGGCTATTAATGCCCTTATTAATCTCAAA 4830	JOURNAL	Submitted (14-OCT-2002) Department of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
Db	50451 TAGTTGTTCAATAGATGAAAGGTGCCTTTGAAAGGGTACTTAAAGAGAACT 50510	REFERENCE	3 (bases 1 to 134544)
Qy	4831 TCCATTATTGCTTCTCTTAATCAGTACCTTGAAGAAACCATGTTGTTAT 4890	AUTHORS	Shailly,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
Db	50511 CTATGTTGAGAACCTAAAGGGTTATGATCTCTCCACATCATGTTAGATAT 50570	TITLE	Direct Submission
Qy	4891 GAACCGCTCTGCTATGGTTAAACAGTCGGTTGGAATGGTATCAACTATCAAAG 4950	JOURNAL	Submitted (13-DEC-2003) Department of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
Db	50571 GAGAAAGCTTGTATGGCTTAAGAGGCACTGGCTGGTATGAGTACACTA 50630	COMMENT	On Dec 13, 2003 this sequence version replaced gi:39752708.
Qy	4951 ACTATTGGAGACATGTTTACTCAAGTTTACACATGCTTTATTCAGTCA 5010	-----	-----
Db	50631 GTTATTTGGTCACGGTACATGAGAGGTACATACAGACCTTT--GAAA 50687	-----	-----
Qy	5011 ATATGAAAGGAGCTACGTAATATTTAGGTTATGTGATGATTTATGGTG 5070	FEATURES	-----
Db	50588 AGATGACATGGAGACTGATGATGCTCATATATATGTTGATGACATGTTTGG 50747	source	-----
Qy	5071 AAGTCACAAAGTTATGATAATTGTTGGATCATGGAGTCATTTGAGTAA 5130	ORIGIN	-----
Db	50748 AATGTCAACACAGATGGTCACATTAATCAGCAATCTGAGTTGATGAT 50807	-----	-----
Qy	5131 AGTGTGTTGCTGAAATTCATTCAATTCTGGATGATTGTAACACGAAATTGTTA 5190	Query Match	1.6%; Score 102.7; DB 2; Length 134544;
Db	50808 CTTAGTCGGTGGTCACTTCTGCTGCTCATGTTAAACCAATGGAGTACTAT 50867	Best Local Similarity	49.5%; Pred. No. 4.2e-07; Mismatches 263; Conservative 0; Mismatches 268; Indels 0; Gaps 0;
Qy	5191 TATTTATCTCAGAAATTCCTCAAGAAATTCTCAAGAAATTCTCAAGAACTGACTC 5250	Matches	263; Conservative 0; Mismatches 268; Indels 0; Gaps 0;
Db	50868 CTCTCATCTCAAGAAATATGCCAGAGCTAGTGAAGAAATTGGTCAAA 50927	ACCESSION	AC135396
Qy	5251 ATATGGGAAACATACCTGGATTGCAAGACAAATATAAAGGTCAAA 5303	VERSION	AC135396.25 GI:39841085
Db	50928 TGGCATAAGAAACTCCAACTGCACTTGAACCTGAAAGGAA 50980	SOURCE	HTG: HTGS PHASE1: HTGS DRAFT,
Qy	4844 TTCTCTCTAAATCAGTACCTTGTGAAAGAACCATGTCGTTGTTGACGTC 4903	ORGANISM	Medicago truncatula (barrel medic)

Db	30251	CCACCAAGTTGAAATTGCTTAGGAGGAATATGGTGTATAACTGCATAAGCACTA	30192
Qy	4904	TATGGGTTAAACACAGCGGGTTGATGGATCACACTCAAAAGGATATCGGAAGC	4463
Db	30191	TATGGGTTAAACACAGCGGTCTTCAAGGATGGACAGAAGAAATTGATCAAGTTATTCAA	30132
Qy	4964	ATTGGTTTACTCAGGTTTACACATGAGGTTATTTCATGAAATGAGGGA	5023
Db	30131	ATTAGGCTTTAAAGTGTCACTGAGTGTGTTGATAGTGTAGCTAGGAAAGCTAGTGTGA	60072
Qy	5024	TCTAGTATATATTAGGTTATGTCAGATATCTATGGTGGAAAGTCACAAA	5083
Db	30071	CGAACCTGTTAATTCATTCGTCATGATGACTCTACTCTACTAGTGTAGTCACATCA	30012
Qy	5084	GTATGGATAATTGGTGTGATCAATGGAGACATCATTTGAGTTAAGTTGGTGA	5143
Db	30011	GAGATTAAGAGGTTAAGAGAACCTGAGCTGGAGTTGAGACTGACTGTTGGAG	29952
Qy	5144	ATATCAATTATCTTGTATGAAATCGTAAACCGAACCTGTTATTTATCTCA	5203
Db	29951	CTTCATCTCTCTAGGAATGAAATTGGATAGTGGATAGTGTGATGATCA	29892
Qy	5204	GAATAATTCTCAAGAATTCTTAAAGATTCACACTAGTGTGATCATAT	5254
Db	29891	AAAATAACATGGTGAATTGTGATAAATTGAAATGAGAGCTGCAAT	29841

Search completed: April 11, 2004, 18:29:14
 Job time : 15982 secs

This Page Blank (uspto)

PT readthrough_protein
 PT /note= "The gag-pol readthrough protein is the result of
 PT stop codon suppression mediated by pseudoknot formation
 PT in the mRNA."
 PT /codon= {seq:"ctg", aa:Ser}
 PT /note= "12 serine residues in the gag-pol readthrough
 PT protein are encoded by this non-standard Ser codon"
 PT /transl_except= {pos:1367..1372, aa:Glu}
 PT 498..1372
 PT /*tag= h
 PT CDS
 PT /product= "Gag (group antigen) protein"
 PT /note= "5 Serine residues in the gag protein are encoded
 PT by this non-standard Ser codon"
 PT /note= "7 serine residues in the pol protein are encoded
 PT by this non-standard Ser codon"
 PT /tag= k
 PT /product= "Pol (polyprotein), comprising aspartate
 PT protease, integrase, reverse transcriptase (RT) and
 PT RNaseH"
 PT /codon= {seq:"ctg", aa:Ser}
 PT /note= "The RNA corresponding to this region forms a
 PT pseudoknot, allowing gag ORF stop codon suppression and
 PT translation of the gag-pol readthrough protein"
 PT 3455..3465
 PT /*tag= l
 PT /note= "Polypurine tract 2 (PPT2)"
 PT complement (3455..3465)
 PT /*tag= m
 PT /note= "This site in the corresponding RNA is a primer
 PT binding site for reverse transcriptase-mediated plus-
 strand DNA synthesis"
 PT 6136..6146
 PT /*tag= n
 PT /note= "Polypurine tract 1 (PPT1)"
 PT complement (6136..6146)
 PT /*tag= o
 PT /note= "This site in the corresponding RNA is a primer
 PT binding site for reverse transcriptase-mediated plus-
 strand DNA synthesis"
 PT 6147..6126
 PT /*tag= p
 PT /note= "imperfect 6 bp repeat"
 PT 6266..6271
 PT /*tag= r
 PT /standard_name= "TATA box"
 PT 6331..6336
 PT /*tag= s
 PT /standard_name= "TATA box"
 PT 6347..6351
 PT /*tag= t
 PT 6421..6426
 PT /*tag= u
 PT repeat_unit
 PT /*tag= v
 PT repeat_unit
 PT /*tag= w
 PT /note= "Imperfect 6 bp repeat"
 XX WO200026397-A1.
 XX 11-MAY-2000.
 XX 01-NOV-1999; 99WO-NZ000179.
 XX 30-OCT-1998; 98CA-02249046.
 PR 30-OCT-1998; 98US-0106342P.
 XX (JANCS) JANSSEN PHARM NV.
 XX

PI Luysten WHML, De Backer MD, Nelissen BJM, Poulter RTM;
 XX WPI: 2000-365640/31.
 DR P-PSDB; AAB03126, AAB03127, AAB03128.
 PT Novel retrotransposon expression vectors useful for expressing an
 PT antigen, epitope or therapeutic agent, or detecting genes or the presence
 PT of Candida in a sample.
 XX
 PS Claim 12; Fig 2B; 204pp; English.
 XX
 CC The invention relates to novel retrotransposons from the yeast *Candida*
 CC albicans which have a copy number of 40-150, preferably 50-100 copies per
 CC genome. In particular, the invention relates to the integrated form of
 CC this retrotransposon, designated TCA2, and to the novel *C. albicans*
 CC retrotransposons 1-28, each was initially isolated from *C. albicans*
 CC KOG1042 and has a copy number of 50-100 copies per cell. It comprises
 CC identical 280 bp long terminal repeats (LTRs) and two open reading frames
 CC (ORFs). The first ORF encodes a gag (group antigen) protein, and the
 CC second ORF encodes a polyprotein (pol) consisting of an aspartate
 CC protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and
 CC pol ORFs of PC1 are in the same reading frame, separated only by a
 CC termination codon (TGA). Translation of the pol ORF occurs through the
 CC occasional readthrough suppression of the stop codon, which is mediated
 CC by the formation of a pseudoknot within the gag-pol mRNA. The
 CC retrotransposons of the invention can be used as vectors for *in vitro* or
 CC *in vivo* transformation and expression. They can thus be used for the
 CC delivery and expression of a therapeutic, immunological or immunogenic
 CC molecule (e.g., an antigen) and may also be used for eliciting an
 CC immunological response in a host organism. They are therefore useful in
 CC genetic vaccine compositions and for gene therapy, particularly where the
 CC use of retroviral vectors is unsafe or undesirable. Additionally, the
 CC retrotransposons may be used to generate transgenic animals, to detect
 CC the presence of Candida in a sample, to detect and disrupt genes, and to
 CC assign functions to nucleotide sequences. The present sequence represents
 CC the *C. albicans* pc1 retrotransposon. (Updated on 15-SEP-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 6426 BP; 2228 A; 1032 C; 1258 G; 1908 T; 0 U; 0 Other;
 Query Match 100.0%; Score 6424.4; DB 3; Length 6426;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 6425; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 TGTGTGTTGTCACATTGTGAGAACTGATCACTGAAAGTATGATGA 60
 DB 1 TGTGTGTTGTCACATTGTGAGAACTGATCACTGAAAGTATGATGA 60
 OY 61 GATGTGAAATTTCATCCACACAGGATGAGCAACTATATGGTAGT 120
 DB 61 GATGTGAAATTTCATCCACACAGGATGAGCAACTATATGGTAGT 120
 OY 121 ATTAATAGGTATGATAACCTCCGAATTCACAGGATGAGGGAGTT 180
 DB 121 ATTAATAGGTATGATAACCTCCGAATTCACAGGATGAGGGAGTT 180
 OY 181 TCAATATATCTGTGATAATGTTCTATTCACATACACACTAGGGTT 240
 DB 181 TCAATATATCTGTGATAATGTTCTATTCACATACACACTAGGGTT 240
 OY 241 ACACGCTCAATCTGAGAAAGTTATTCACATACACACTAGGGTT 300
 DB 241 ACACGCTCAATCTGAGAAAGTTATTCACATACACACTAGGGTT 300
 OY 301 ATTCATTCGTCACAATTAGGTGTTATAATTCTGCTCTGAGTTGATT 360
 DB 301 ATTCATTCGTCACAATTAGGTGTTATAATTCTGCTCTGAGTTGATT 360
 OY 361 GATAGTTGAGGTGAGGACAGATTCACAGATGAGTGGTCGCAAGAATGATGA 420
 DB 361 GATAGTTGAGGTGAGGACAGATTCACAGATGAGTGGTCGCAAGAATGATGA 420

Qy	421	TAACGAAAGGGAGGTCACTGAAAGTGTGACCAAGCTAATGCTATTAGTAAAGGTGATGAA	480	Db	1501	TAACPAGATCTAGTGTAGTGTACTGGTGCACAACTCTGTGIGACAAATAGAT	1560
Db	421	TAACGAAAGGGAGGTCACTGAAAGTGTGACCAAGCTAATGCTATTAGTAAAGGTGATGAA	480	Qy	1561	ATTCGTGATGTTAAGGCGACAATGAAAGTTCTGTCGTCGTTGCTGTTGCTACATAGA	1620
Qy	481	ACATATCAGGCTAGATTCAATATGTTTCAATAATTATGACTTACCTAGTTGCG	540	Db	1561	ATTGTGATGTTAAGGCGACAATGAAAGTTCTGTCGTCGTTGCTGTCACATAGA	1620
Db	481	ACATATCAGGCTAGATTCAATATGTTTCAATAATTATGACTTACCTAGTTGCG	540	Qy	1621	AGGAGATGTAATGGTGTCTTATACAGAGCTCGTATTGCTCATAGTGTAGAA	1680
Qy	541	CTCGGTATAGAAGAAGGGATAATGAAAGAGATTAATTCAGGTG	600	Db	1621	AGCAGATGTTATGGTGTCTTATACAGAGCTCGTATTGCTCATAGTGTAGAA	1680
Db	541	CTCGGTATAGAAGAAGGGATAATGAAAGAGATTAATTCAGGTG	600	Qy	1681	TACATGATTCAGAGAGTCTTATCTGAGTCGTTGCTGAGTTGAAACAAATGAGAGC	1740
Qy	601	TTACCCGATTTGGAAATTGCTGACTATATCCAAAGATAATTCCAGTTAA	660	Db	1681	TACATGATTCAGAGAGTCTTATCTGAGTCGTTGCTGAGTTGAAACAAATGAGAGC	1740
Db	601	TTACCCGATTTGGAAATTGCTGACTATATCCAAAGATAATTCCAGTTAA	660	Qy	1741	AGGATTAAATGTTCTTACTAAGTGTGTTACGTCATT	1800
Qy	661	AAAGGTAGAAAGTATTATTACTGTTGTTACGTCATT	720	Db	1741	AGGATTAAATGTTCTTACTAAGTGTGTTACGTCATT	1800
Db	661	AAAGGTAGAAAGTATTATTACTGTTGTTACGTCATT	720	Qy	1801	TACATGATTCAGAGAGTCTTATCTGAGTCGTTGCTGAGTTGAAACAAATGAGAGC	1860
Qy	721	TGATAGGTTGAGTATGATGATCAGTAACTACCAAGACTGCAAGGAGAA	780	Db	1801	TACATGATTCAGAGAGTCTTATCTGAGTCGTTGCTGAGTTGAAACAAATGAGAGC	1860
Db	721	TGATAGGTTGAGTATGATCAGTAACTACCAAGACTGCAAGGAGAA	780	Qy	1861	AGAATCTTGTGTTAGATGTTGATGTTGATGTTGAGTGTGTTGAGAATGTC	1920
Qy	781	CTTATCAAGGCTTATGTCGTTACATCTAACATTAAGATTAACCTACTA	840	Db	1861	AGAATCTTGTGTTAGATGTTGATGTTGATGTTGAGTGTGTTGAGAATGTC	1920
Db	781	CTTATCAAGGCTTATGTCGTTACATCTAACATTAAGATTAACCTACTA	840	Qy	1921	CCAGATGACAAAGATAATCACTGTTGAACTGCTAGATTCAGATTCAGATGAA	1980
Qy	841	GATCTCTGTGAGTTGAAAGTGTGACATGAGTTAGTGGTGTCTAACCTCCATA	900	Db	1921	CCAGATGACAAAGATAATCACTGTTGAACTGCTAGATTCAGATTCAGATGAA	1980
Db	841	GATCTCTGTGAGTTGAAAGTGTGACATGAGTTAGTGGTGTCTAACCTCCATA	900	Qy	1981	TAGTCTGAGATTAATTCAGTGTGTTGAGTGTGTTGAGTGTGTTGAGTGAATT	2040
Qy	901	TTCGTCGAGGTGGAAAGAAACTGAGGAATACCTACACACTTCACAACTGTC	960	Db	1981	TAGTCTGAGATTAATTCAGTGTGTTGAGTGTGTTGAGTGTGTTGAGTGAATT	2040
Db	901	TTCGTCGAGGTGGAAAGAAACTGAGGAATACCTACACACTTCACAACTGTC	960	Qy	2041	CCCATATGGAGTGAACATGCTACCAACTCGAGAGATAAGCATGTTAATTC	2100
Qy	961	TGAGTAGTGTGTTAGTCTCAATCTACAGTGTGTTACTACATGTTGAA	1020	Db	2041	CCCATATGGAGTGAACATGCTACCAACTCGAGAGATAAGCATGTTAATTC	2100
Db	961	TGAGTAGTGTGTTAGTCTCAATCTACAGTGTGTTACTACATGTTGAA	1020	Qy	2101	TTTGTGCAACTATGTCATTGAGAAATCTGTTGTTACAAAAATCAGGCT	2160
Qy	1021	GAATCAGAGGCTTGTAGTTGTTGATATTCTGCTCATACTCAATGAGTGG	1080	Db	2101	TTTGTGCAACTATGTCATTGAGAAATCTGTTGTTACAAAAATCAGGCT	2160
Db	1021	GAATCAGAGGCTTGTAGTTGTTGATATTCTGCTCATACTCAATGAGTGG	1080	Qy	2151	CGFACTTCACACTTCACAAAGAGACTCTCAAAAGATGTCGTTGAGTGTCTT	2220
Qy	1081	TCAGCTGACATGATCATAGTACTCCAGTTCTTACCATAGTCAGAAATG	1140	Db	2151	CGFACTTCACACTTCACAAAGAGACTCTCAAAAGATGTCGTTGAGTGTCTT	2220
Db	1081	TCAGCTGACATGATCATAGTACTCCAGTTCTTACCATAGTCAGAAATG	1140	Qy	2221	ATCGAACCCACAGAGAGCACATCATCAGTCAGAAAGACCTCTGAGAGA	2280
Qy	1141	TGATCATCTCTATTATGCTAGTTGAGTGTCTTACCATAGTCAGAAATG	1200	Db	2221	ATCGAACCCACAGAGAGCACATCATCAGTCAGAAAGACCTCTGAGAGA	2280
Db	1141	TGATCATCTCTATTATGCTAGTTGAGTGTCTTACCATAGTCAGAAATG	1200	Qy	2281	TAAGAGATTCATGTCATCTCGCTTCACTTGTGAAATTAACAGGGATT	2340
Qy	1201	TATGTTAGTTTACCGACACAGAAACCAAGAGAGACTCACTGCTGAA	1260	Db	2281	TAAGAGATTCATGTCATCTCGCTTCACTTGTGAAATTAACAGGGATT	2340
Db	1201	TATGTTAGTTTACCGACACAGAAACCAAGAGAGACTCACTGCTGAA	1260	Qy	2341	AACTCTGTTATGATGACATACGGTTACATGTCGTTCACTTGTGAAATTA	2400
Qy	1261	ACAAATCTAAAGAGAACTGAAATCAAAAGAGAAATAGAACATCCAAATCGA	1320	Db	2341	AACTCTGTTATGATGACATACGGTTACATGTCGTTCACTTGTGAAATTA	2400
Db	1261	ACAAATCTAAAGAGAACTGAAATCAAAAGAGAAATAGAACATCCAAATCGA	1320	Qy	2401	GGAAAGGATCTTAACTCAGATAATCTGAGTGTCTGAGTTCACACTCTG	2460
Qy	1321	TAACGATAAAGAGAAACCTACTGGATGAAACAG	1380	Db	2401	GGAAAGGATCTTAACTCAGATAATCTGAGTGTCTGAGTTCACACTCTG	2460
Db	1321	TAACGATAAAGAGAAACCTACTGGATGAAACAG	1380	Qy	2461	GGCAATTGGAGGAGCTATACGGCTATCTGCTGAGTGTCTGAGTTCAC	2520
Qy	1381	TGCTGCTCTTAACTCTGAGTATGTCGAAACAGCTTCCAGT	1440	Db	2461	GGCAATTGGAGGAGCTATACGGCTATCTGCTGAGTGTCTGAGTTCAC	2520
Db	1381	TGCTGCTCTTAACTCTGAGTATGTCGAAACAGCTTCCAGT	1440	Qy	2521	CGGATTTGGAGGAGCTATACGGCTATCTGCTGAGTGTCTGAGTTCAC	2580
Qy	1441	AGAAAATTCTTCTTAATGCTCTTGGACGAAATGATTTAAGGTTAGGT	1500	Db	2521	CGGATTTGGAGGAGCTATACGGCTATCTGCTGAGTGTCTGAGTTCAC	2580
Db	1441	AGAAAATTCTTCTTAATGCTCTTGGACGAAATGATTTAAGGTTAGGT	1500	Qy	2581	TGTTAATTAATGATTAACAGATTTACAGGATGTTGACACTGTC	2640
Qy	1501	TRACAAGTATGATTAACAGATTTACAGGATGTTGACACTGTC	1560				

Db	2581	TGTTAATAATTGATTACACAGATTACAGGTCTGCTGACACTTGGCCACAAAT	QY	3721	ATTCCGGGGGTGATGATTCTGCTGATCACTAAATGTTGAGCAATCTATGT
QY	2641	ACTCAAGTGTGATTATATGATTCATAATCTTACATGATCAACACACTCCAGG	Db	3721	ATTCCGGGGGTGATGATTCTGCTGATCACTAAATGTTGAGCAATCTATGT
Db	2641	ACTCAAGTGTGATTATATGATTCATAATCTTACATGATCAACACACTCCAGG	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	2701	TGTTACTCAGGGCAACCCCTATGGTGTCTATTAACTAATAGGAGGAATT	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	2701	TGTTACTCAGGGCAACCCCTATGGTGTCTATTAACTAATAGGAGGAATT	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	2761	CTACCGTTCTTGCATGATGTTGCTACATTAGTATGCTAAGTGAAGAA	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	2761	CTACCGTTCTTGCATGATGTTGCTACATTAGTATGCTAAGTGAAGAA	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	2821	CGGTTAGGAGTACATCACTAAAGGA	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	2821	CGGTTAGGAGTACATCACTAAAGGA	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	2881	CTACGTAGCAGTTAGTATAGTGTGCTAALAAATGCGGTGATATT	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	2881	CTACGTAGCAGTTAGTATAGTGTGCTAALAAATGCGGTGATATT	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	2941	CCTTAGCCTATGTCGATATGGATATCTTACATGGGTCTGATGG	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	2941	CCTTAGCCTATGTCGATATGGATATCTTACATGGGTCTGATGG	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3001	CTTATCACTACACCTATGTCGATATGGGTCTATGGTGAAGGATCAGGAA	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3001	CTTATCACTACACCTATGTCGATATGGGTCTATGGTGAAGGATCAGGAA	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3061	ACTGGGCGCTAGTAGGAGGACATATGGAAGTGAATGACATACAA	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3061	ACTGGGCGCTAGTAGGAGGACATATGGAAGTGAATGACATACAA	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3121	TGACGTGATGACATGCCAGAGGTCAATCAGTCACTTAA	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3121	TGACGTGATGACATGCCAGAGGTCAATCAGTCACTTAA	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3181	TAGGGTACAGTTAACGATATGATGATGACAGTACAGTACATACAA	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3181	TAGGGTACAGTTAACGATATGATGACAGTACAGTACATACAA	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3241	GTATCCGATATTTCTAACCTTTCACTACTGTGAGATCACATATGTC	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3241	GTATCCGATATTTCTAACCTTTCACTACTGTGAGATCACATATGTC	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3301	CGAAGTAAATGGTGAAGAATCTAACCCAGTCTCCAGGAGTACACCTGGGATA	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3301	CGAAGTAAATGGTGAAGAATCTAACCCAGTCTCCAGGAGTACACCTGGGATA	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3361	TGGGTCTTAACCTCTACTGTGAGGACTTCAGTAATGGGAACTTAACG	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3361	TGGGTCTTAACCTCTACTGTGAGGACTTCAGTAATGGGAACTTAACG	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3421	GCCTATTCAAAACCAAAAGGCCCTTCATCCAGGGAGGACATAAACGCT	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3421	GCCTATTCAAAACCAAAAGGCCCTTCATCCAGGGAGGACATAAACGCT	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3481	GGATCTGTGAGGTTGACTACAGTACCTGTTGAGGAGCTTACAGT	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3481	GGATCTGTGAGGTTGACTACAGTACCTGTTGAGGAGCTTACAGT	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3541	GGATCTGTGAGGTTGACTACAGTACCTGTTGAGGAGCTTACAGT	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3481	GGATCTGTGAGGTTGACTACAGTACCTGTTGAGGAGCTTACAGT	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3601	AGAGATGGCATACTCTGTTACCCGAAACTAGTGCAAGAGTGAAGAAC	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3601	AGAGATGGCATACTCTGTTACCCGAAACTAGTGCAAGAGTGAAGAAC	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
QY	3661	CCTGTTACTGGCTAATAGAATCGGTAACGAAANATAGTGGGAAATTTC	Db	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT
Db	3661	CCTGTTACTGGCTAATAGAATCGGTAACGAAANATAGTGGGAAATTTC	QY	3781	TGAAATGAGGATGCTGCTGATCAGCTCAATCAGAGGAGTTCTCAGAGGAGAAT

QY	4801	GTATCTAAATGCCCTTAACTCAATCCAAATTATGCTTCCCTCCAAATCAGT	4860	Db	5881	TCACCCACATGCCAAATAGAATCTGCTATAGTTCTAGCCAATGACAATGAA	5940
Db	4801	GATCCTAAATGCCCTTAACTCAATCCAAATTATGCTTCCCTCCAAATCAGT	4860	QY	5941	AGTATTCATATCCTATGAGAACATGATAATACCGCGATTTGACTAATGG	6000
QY	4861	ACCTTGAGAAACCATTTGTTTATTCACCTTCTCTATGGGTTAACAGTC	4920	Db	5941	AGTATTCATATCCTATGAGAACATGATAATACCGCGATTTGACTAATGG	6000
Db	4861	ACCTTGAGAAACCATTTGTTTATTCACCTTCTCTATGGGTTAACAGTC	4920	QY	6001	TCTAAGTGTGCTAAATCAGCATGGTGGGTTATGATAAACGGTTRAGCTGA	6060
QY	4921	CGGTGTTGAAATGTTACACTATCAAAAGGTATGGAGACATGGTTTACICAGTC	4980	Db	6001	TCTAAGTGTGCTAAATCAGCATGGTGGGTTATGATAAACGGTTRAGCTGA	6060
Db	4921	GGGTGTTGAAATGTTACACTATCAAAAGGTATGGAGACATGGTTTACICAGTC	4980	QY	6051	AGATAATCCTAAACATGCTACAAATGCAAAAGGATGTTACATCACT	6120
QY	4981	TTTACACAAATGTTTACATGTTACACTATCAAAAGGTATGGAGACATGGTTTACICAGTC	5040	Db	6051	AGTATTAACACTGCTACAAATGCAAAAGGATGTTACATCACT	6120
Db	4981	TTTACACAAATGTTTACATGTTTACATGAAATGAGAGGATGTTACATATTTAGG	5040	QY	6121	ATRATCGATGCTCAATCAGGGAGTTGGTTGCACTATTGTCAGAACG	6180
QY	5041	TTTATGTTGTTGATTCCTATGGTGGAGTTCACAAAGGATTTGTTACATTTG	5100	Db	6121	ATRATCGATGCTCAATCAGGGAGTTGGTTGCACTATTGTCAGAACG	6180
Db	5041	TTTATGTTGTTGATTCCTATGGTGGAGTTCACAAAGGATTTGTTACATTTG	5100	QY	6181	ATCATGAAATGATGGTTATTAGAGATGGAAATTTTCATCACACAGTGT	6240
QY	5101	GGACATTTGAGAGACATTTGAGTAAGTAAAGGTTGGAAATTCAGTC	5160	Db	6181	ATCATGAAATGATGGTTATTAGAGATGGAAATTTTCATCACACAGTGT	6240
Db	5101	GGACATTTGAGAGACATTTGAGTAAGTAAAGGTTGGAAATTCAGTC	5160	QY	6241	GACGAACTAACTATATGTTGAGTAAATAAGGTGATGAAATCCACATCCAGAA	6300
QY	5161	TATGAAATTGTTGATAAACCGATCTGGTATTTTACATCAAAATCTCG	5220	Db	6241	GACGAACTAACTATATGTTGAGTAAATAAGGTGATGAAATCCACATCCAGAA	6300
Db	5161	TATGAAATTGTTGATAAACCGATCTGGTATTTTACATCAAAATCTCG	5220	QY	6301	TATCAACCGAGATAAGGAGGTTCAATATCTGTGAAATAACTCTGTT	6360
QY	5221	ATTACTTAAGATTCAACTAGATGACTGATGTTACATGGGATCTGG	5280	Db	6301	TATCAACCGAGATAAGGAGGTTCAATATCTGTGAAATAACTCTGTT	6360
Db	5221	ATTACTTAAGATTCAACTAGATGACTGATGTTACATGGGATCTGG	5280	QY	6361	ATTCAATATCACACTAGAGCTGATGACCCCTCATTCGGTAAGAAGTTATT	6420
QY	5281	TGACAATATGAAATGGTGGAAATTCTGAAACGTTATCCAGAGATGTTG	5340	Db	6361	ATTCAATATCACACTAGAGCTGACCCCTCATTCGGTAAGAAGTTATT	6420
Db	5281	TGACAATATGAAATGGTGGAAATTCTGAAACGTTATCCAGAGATGTTG	5340	QY	6421	CCATCA 6426	
QY	5341	AAAGTTCCGAGAGACATGCTGACCGTGTGACCTGCTAAAAACTATCCAACTGGT	5400	Db	6421	CCATCA 6426	
Db	5341	AAAGTTCCGAGAGACATGCTGACCGTGTGACCTGCTAAAAACTATCCAACTGGT	5400	QY			
QY	5401	TGGCCCTGCTTTATGGCTGCCAACACAGCGTCAGATATCGGTGTTGAAATTC	5460				
Db	5401	TGGCCCTGCTTTATGGCTGCCAACACAGCGTCAGATATCGGTGTTGAAATTC	5460	RESULT 2			
QY	5461	GTTGGATTCATTAATCTGCAATCCAAATGTCATGGTCAATGTTGAGTATGTTGCT	5520	ID	AAA57336	standard; DNA; 392 BP.	
Db	5461	GTTGGATTCATTAATCTGCAATCCAAATGTCATGGTCAATGTTGAGTATGTTGCT	5520	XX	AAA57336;		
QY	5521	TAGTGTATGAAATAGCTGGATGTTGATGAAATTGGGCAATGGCAGATGTTGCT	5580	AC	AAA57936;		
Db	5521	TAGTGTATGAAATAGCTGGATGTTGATGAAATTGGGCAATGGCAGATGTTGCT	5580	XX	15-SEP-2003 (revised)		
QY	5581	ACACCAAAATCATGTTGATGAAATTGGGCAATGGCAGATGTTGCT	5640	DT	10-OCT-2000 (first entry)		
Db	5581	ACACCAAAATCATGTTGATGAAATTGGGCAATGGCAGATGTTGCT	5640	XX	Candida albicans strain p36 Tca2 retrotransposon 5' region.		
QY	5641	TGAAATCTTATGGGACTTGTGATGAAATTGGGCAATGGCAGATGTTGCT	5700	DE			
Db	5641	TGAAATCTTATGGGACTTGTGATGAAATTGGGCAATGGCAGATGTTGCT	5700	XX	KW	Retrotansposon; pCal, Tca2, Tyl; copia; long terminal repeat; LTR;	
QY	5701	CAAAACAAACGGTCATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5760	XX	KW	gag gene; group antigen; polypeptide; pol; aspartate protease; integrase;	
Db	5701	CAAAACAAACGGTCATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5760	XX	KW	reverse transcriptase; ribonuclease; pseudoknot; readthrough translation;	
QY	5760	CAAAACAAACGGTCATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5760	XX	KW	stop codon suppression; gene delivery; gene therapy vector;	
Db	5760	CAAAACAAACGGTCATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5760	XX	KW	genetic vaccine composition; immunogenic; transgenic animal; ds.	
QY	5761	TACAATGTTGAACTATGCAATGAAACCATTTAATGGTTGGTTGAGTAGG	5820	OS	Candida albicans; strain p36.		
Db	5761	TACAATGTTGAACTATGCAATGAAACCATTTAATGGTTGGTTGAGTAGG	5820	XX	KW		
QY	5821	TTAGATACATGTCATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5880	PP	01-NOV-1999; 99W0-NZ000179.		
Db	5821	TTAGATACATGTCATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5880	XX	W0200026397-A1.		
QY	5881	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX	11-MAY-2000.		
Db	5881	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
Db	5940	TCACCCACATGCCAACGCTGAGCTGCTGAAATGGCTTAATTA	5940	XX			
QY</td							

XX Novel retrotransposon expression vectors useful for expressing an antigen, epitope or therapeutic agent, or detecting genes or the presence of *Candida* in a sample.

XX Example 9; Fig 9; 204pp; English.

CC The invention relates to novel retrotransposons from the yeast *Candida albicans* which have a copy number of 40-150, preferably 50-100 copies per genome. In particular, the invention relates to the novel *C. albicans* Ty1/copia retrotransposon pCAL (AA57920), and to the integrated form of this retrotransposon, designated TCA2, and to the novel *C. albicans* retrotransposons 1-28. pCAL was initially isolated from *C. albicans* hOG142 and has a copy number of 50-100 copies per cell. It comprises identical 280 bp long terminal repeats (LTRs) and two open reading frames (ORFs). The first ORF encodes a gag (group antigen) protein, and the second ORF encodes a polypeptide (pol) consisting of an aspartate protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and pol ORFs of pCAL are in the same reading frame, separated only by a termination codon (TGA). Translation of the pol ORF occurs through the occasional readthrough suppression of the stop codon, which is mediated by the formation of a pseudoknot within the gag-pol mRNA. The retrotransposons of the invention can be used as vectors for *in vitro* or *in vivo* transformation and expression. They can thus be used for the delivery and expression of a therapeutic, immunological or immunogenic molecule (e.g., an antigen) and may also be used for eliciting an immunological response in a host organism. They are therefore useful in generic vaccine compositions and for gene therapy, particularly where the use of retroviral vectors is unsafe or undesirable. Additionally, the presence of *Candida* in a sample, to detect and disrupt genes, and to assign functions to nucleotide sequences. Sequences AA57923-AA57936 represent TCA2 retrotransposon 5' regions from a variety of *C. albicans* strains. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 392 BP; 136 A; 56 C; 72 G; 128 T; 0 U; 0 Other;

XX Query Match 6.1%; Score 392; DB 3; Length 392;

CC Best Local Similarity 100.0%; Pred. No. 7.9e-69; Mismatches 0; Indels 0; Gaps 0; Matches 392; Conservative 0;

QY 1 TGTGTTGTTGCACTATTGTCAGAACTGTCATAAGAATGATGTTATTATGA 60

Db 1 TGTGTTGTTGCACTATTGTCAGAACTGTCATAAGAATGATGTTATTATGA 60

QY 61 GATGGAAATTTCATCACACATCAGGATGATCAGAACTAACTATATGTTAGT 120

Db 61 GATGGAAATTTCATCACACATCAGGATGATCAGAACTAACTATATGTTAGT 120

QY 121 ATAAATAGGTTAGAAATACACACATCCAGAATTCAGGAGATAGGGAGCT 180

Db 121 ATAAATAGGTTAGAAATACACACATCCAGAATTCAGGAGATAGGGAGCT 180

QY 181 TCAATATATCTGAGATAAACTCTGTTAACTACACAACTAGAGCT 240

Db 181 TCAATATATCTGAGATAAACTCTGTTAACTACACAACTAGAGCT 240

QY 241 ACAGGTCATACTCAGGAAAGTTATTCATCAAGATAGGAGATGAT 300

Db 241 ACAGGTCATACTCAGGAAAGTTATTCATCAAGATAGGAGATGAT 300

QY 301 AACATATTGTCGCAATTAGCGTTATAATTCACTGCTCAGATGTTATTGATT 360

Db 301 AACATATTGTCGCAATTAGCGTTATAATTCACTGCTCAGATGTTATTGATT 360

XX Sequence 392 BP; 136 A; 56 C; 72 G; 128 T; 0 U; 0 Other;

XX Query Match 6.1%; Score 388.8%; DB 3; Length 392;

CC Best Local Similarity 99.5%; Pred. No. 7.9e-69; Mismatches 2; Indels 0; Gaps 0; Matches 390; Conservative 0;

QY 1 TGTGTTGTTGCACTATTGTCAGAACTGTCATAAGAATGATGTTATTATGA 60

RESULT 3
AA57928 standard; DNA; 392 BP.

XX AC AA57928;

XX PT 15-SEP-2003 (revised)

PT DT 10-OCT-2000 (first entry)

XX DE Candida albicans strain ATC-2 TCA2 retrotransposon 5' region.

XX KW Retrotransposon; pCAL; TCA2; Ty1; copia; long terminal repeat; LTR; gag; group antigen; polypeptide; pol; aspartate protease; integrase; reverse transcriptase; RNaseH; pseudoknot; readthrough translation; stop codon suppression; gene delivery; gene therapy vector; genetic vaccine composition; immunogenic; transgenic animal; ds.

XX OS Candida albicans; strain ATC-2.

XX PN WO20026397-A1.

XX PD 11-MAY-2000.

XX PF 01-NOV-1999; 99WO-NZ000179.

XX PR 30-OCT-1998; 98CA-02249046.

XX PR 30-OCT-1998; 98US-0106342P.

XX PA (JANCO) JANSSEN PHARM NV.

XX PT Luyten WHML, De Backer MD, Nelissen BJM, Poulter RTM;

XX DR WPI; 2000-365640/31.

XX PT Novel retrotransposon expression vectors useful for expressing an antigen, epitope or therapeutic agent, or detecting genes or the presence of *Candida* in a sample.

XX PT Novel retrotransposon expression vectors useful for expressing an antigen, epitope or therapeutic agent, or detecting genes or the presence of *Candida* in a sample.

XX PS Example 9; Fig 9; 204pp; English.

CC The invention relates to novel retrotransposons from the yeast *Candida albicans* which have a copy number of 40-150, preferably 50-100 copies per genome. In particular, the invention relates to the novel *C. albicans* Ty1/copia retrotransposon pCAL (AA57920), and to the integrated form of this retrotransposon, designated TCA2, and to the novel *C. albicans* hOG142 and has a copy number of 50-100 copies per cell. It comprises identical 280 bp long terminal repeats (LTRs) and two open reading frames (ORFs). The first ORF encodes a gag (group antigen) protein, and the second ORF encodes a polypeptide (pol) consisting of an aspartate protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and pol ORFs of pCAL are in the same reading frame, separated only by a termination codon (TGA). Translation of the pol ORF occurs through the occasional readthrough suppression of the stop codon, which is mediated by the formation of a pseudoknot within the gag-pol mRNA. The retrotransposons of the invention can be used as vectors for *in vitro* or *in vivo* transformation and expression. They can thus be used for the delivery and expression of a therapeutic, immunological or immunogenic molecule (e.g., an antigen) and may also be used for eliciting an immunological response in a host organism. They are therefore useful in generic vaccine compositions and for gene therapy, particularly where the use of retroviral vectors is unsafe or undesirable. Additionally, the retrotransposons may be used to generate transgenic animals, to detect the presence of *Candida* in a sample, to detect and disrupt genes, and to assign functions to nucleotide sequences. Sequences AA57923-AA57936 represent TCA2 retrotransposon 5' regions from a variety of *C. albicans* strains. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 392 BP; 136 A; 56 C; 72 G; 128 T; 0 U; 0 Other;

XX Query Match 6.1%; Score 388.8%; DB 3; Length 392;

CC Best Local Similarity 99.5%; Pred. No. 7.9e-69; Mismatches 2; Indels 0; Gaps 0; Matches 390; Conservative 0;

QY 1 TGTGTTGTTGCACTATTGTCAGAACTGTCATAAGAATGATGTTATTATGA 60

```

Db 1 TGTGGTTGCACTATTGTGTCAGAAATGATCAATGAAATGATGTTTATGA 60
QY 61 GATGGAAATTTCATACACATCAGTGACAGACTAACTATTTGTT 120
Db 61 GATGGAAATTTCATACACATCAGTGACAGACTAACTATTTGTT 120
QY 121 ATAATTAAGGTTATGAAATACACACCTCCAGTATCAAGAGATAGGGAGT 180
Db 121 ATAATTAAGGTTATGAAATACACACCTCCAGTATCAAGAGATAGGGAGT 180
QY 181 TCAATATATCTTGAACTTACACACTAGCT 240
Db 181 TCAATATATCTTGAACTTACACACTAGCT 240
QY 241 ACAGCTCAATCTAGTAAGAACTTATTCATAGTAACTGAGGAGGT 300
Db 241 ACAGCTCAATCTAGTAAGAACTTATTCATAGTAACTGAGGAGGT 300
QY 301 ATCATTTGCGCCAAATTAGCTGTATAATTAGTCCTCAGTTGTTATTGATT 360
Db 301 ATCATTTGCGCCAAATTAGCTGTATAATTAGTCCTCAGTTGTTATTGATT 360
QY 361 GATAGTTGAGTTGAAGGTACAGAATTTC 392
Db 361 GATAGTTGAGTTGAAGGTACAGAATTTC 392

```

RESULT 4

```

AAA57931
ID AAA57931 standard; DNA; 392 BP.
XX
AC AAA57931;
XX
DT 15-SEP-2003 (revised)
DT 10-OCT-2000 (first entry)
XX
DE Candida albicans strain Fl6-1 TCA2 retrotransposon 5' region.
XX
KW Retrotansposon; pCal; Tca2; Ty1; copia; long terminal repeat; LTR;
KW gag gene; group antigen; polyprotein; pol; aspartate protease; integrase;
KW reverse transcriptase; RNaseH; pseudoknot; readthrough translation;
KW stop codon suppression; gene delivery; gene therapy vector;
KW genetic vaccine composition; immunogenic; transgenic animal; ds.
OS Candida albicans; strain Fl6-1.
XX
PN WO20026397-A1.
XX
PD 11-MAY-2000.
XX
PP 01-NOV-1999; 99WO-NZ000179.
XX
PR 30-OCT-1998; 99CA-02249046.
PR 30-OCT-1998; 98US-0106342P.
XX
PA (JANCS ) JANSSEN PHARM NV.
XX
PI Luyten WHM, De Backer MD, Nelissen BJM, Poulter RTM;
XX
DR WPI; 2000-365640/31.
XX
PT Novel retrotransposon expression vectors useful for expressing an
PT antigen, epitope or therapeutic agent, or detecting genes or the presence
PT of Candida in a sample.
XX
PS Example 9; Fig 9; 204pp; English.

```

The invention relates to novel retrotransposons from the yeast *Candida* genus. In particular, the invention relates to the novel *C. albicans* retrotransposon pCal (AAA57931), and to the integrated form of this retrotransposon, designated TCA2, and to the novel *C. albicans* retrotransposons 1-28. pCal was initially isolated from *C. albicans*

h01042 and has a copy number of 50-100 copies per cell. It comprises identical 280 bp long terminal repeats (LTRs) and two open reading frames (ORFs). The first ORF encodes a gag (group antigen) protein, and the second ORF encodes a Pol/protein (pol) consisting of an aspartate protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and pol ORFs of pCal are in the same reading frame, separated only by a termination codon (TGA). Translation of the pol ORF occurs through the occasional readthrough suppression of the stop codon, which is mediated by the formation of a pseudoknot within the gag-pol mRNA. The retrotransposons of the invention can be used as vectors for in vitro or in vivo transformation and expression. They can thus be used for the delivery and expression of a therapeutic, immunological or immunogenic molecule (e.g., an antigen) and may also be used for eliciting an immunological response in a host organism. They are therefore useful in genetic vaccine compositions and for gene therapy, particularly where the use of retroviral vectors is unsafe or undesirable. Additionally, the retrotransposons may be used to generate transgenic animals to detect the presence of *Candida* in a sample, to detect and disrupt genes, and to assign functions to nucleotide sequences. Sequences AAA57931-A57936 represents TCA2 retrotransposon 5' regions from a variety of *C. albicans* strains. (Updated on 15-SEP-2003 to standardise OS field)

Sequence 392 BP; 136 A; 56 C; 72 G; 128 T; 0 U; 0 Other;

Query Match 6.1%; Score 388.8; DB 3; Length 392;

Best Local Similarity 99.5%; Pred. No. 7.9e-69; Mismatches 0; Index 0; Gaps 0;

Matches 390; Conservative 0; Mismatches 2; Index 0; Gaps 0;

Matches 390; Conservative 0; Mismatches 2; Index 0; Gaps 0;

QY 1 TGTGGTTGCACTATTGTGAGAACTGATCATGAAATGATGTTTATGA 60
Do 1 TGTGGTTGCACTATTGTGAGAACTGATCATGAAATGATGTTTATGA 60
QY 61 GATGGAAATTTCATACACACTAGGATAGCAAGACTAACTATTTGTT 120
Do 61 GATGGAAATTTCATACACACTAGGATAGCAAGACTAACTATTTGTT 120
QY 121 ATAATTAAGGTTATGAAATACACACCTCCAGAATTCAGGAGATAGGGAGTT 180
Do 121 ATAATTAAGGTTATGAAATACACACCTCCAGAATTCAGGAGATAGGGAGTT 180
QY 181 TCAATATATCTTGAACTTACACACTAGCT 240
Do 181 TCAATATATCTTGAACTTACACACTAGCT 240
QY 241 ACAGCTCAATCTAGTAAGAACTTATTCATCAGTTGAGTATGTT 300
Do 241 ACAGCTCAATCTAGTAAGAACTTATTCATCAGTTGAGTATGTT 300
QY 301 ATCATTTGCGCCAAATTAGCTGTATAATTAGTCCTCAGTTGTTATTGATT 360
Do 301 ATCATTTGCGCCAAATTAGCTGTATAATTAGTCCTCAGTTGTTATTGATT 360
QY 361 GATAGTTGAGTTGAAGGTACAGAATTTC 392
Do 361 GATAGTTGAGTTGAAGGTACAGAATTTC 392

CC identical 280 bp long terminal repeats (LTRs) and two open reading frames (ORFs). The first ORF encodes a Pol/protein (pol) consisting of an aspartate protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and pol ORFs of pCal are in the same reading frame, separated only by a termination codon (TGA). Translation of the pol ORF occurs through the occasional readthrough suppression of the stop codon, which is mediated by the formation of a pseudoknot within the gag-pol mRNA. The retrotransposons of the invention can be used as vectors for in vitro or in vivo transformation and expression. They can thus be used for the delivery and expression of a therapeutic, immunological or immunogenic molecule (e.g., an antigen) and may also be used for eliciting an immunological response in a host organism. They are therefore useful in genetic vaccine compositions and for gene therapy, particularly where the use of retroviral vectors is unsafe or undesirable. Additionally, the retrotransposons may be used to generate transgenic animals to detect the presence of *Candida* in a sample, to detect and disrupt genes, and to assign functions to nucleotide sequences. Sequences AAA57931-A57936 represents TCA2 retrotransposon 5' regions from a variety of *C. albicans* strains. (Updated on 15-SEP-2003 to standardise OS field)

Sequence 392 BP; 136 A; 56 C; 72 G; 128 T; 0 U; 0 Other;

Query Match 6.1%; Score 388.8; DB 3; Length 392;

Best Local Similarity 99.5%; Pred. No. 7.9e-69; Mismatches 0; Index 0; Gaps 0;

Matches 390; Conservative 0; Mismatches 2; Index 0; Gaps 0;

Matches 390; Conservative 0; Mismatches 2; Index 0; Gaps 0;

QY 1 TGTGGTTGCACTATTGTGAGAACTGATCATGAAATGATGTTTATGA 60
Do 1 TGTGGTTGCACTATTGTGAGAACTGATCATGAAATGATGTTTATGA 60
QY 61 GATGGAAATTTCATACACACTAGGATAGCAAGACTAACTATTTGTT 120
Do 61 GATGGAAATTTCATACACACTAGGATAGCAAGACTAACTATTTGTT 120
QY 121 ATAATTAAGGTTATGAAATACACACCTCCAGAATTCAGGAGATAGGGAGTT 180
Do 121 ATAATTAAGGTTATGAAATACACACCTCCAGAATTCAGGAGATAGGGAGTT 180
QY 181 TCAATATATCTTGAACTTACACACTAGCT 240
Do 181 TCAATATATCTTGAACTTACACACTAGCT 240
QY 241 ACAGCTCAATCTAGTAAGAACTTATTCATCAGTTGAGTATGTT 300
Do 241 ACAGCTCAATCTAGTAAGAACTTATTCATCAGTTGAGTATGTT 300
QY 301 ATCATTTGCGCCAAATTAGCTGTATAATTAGTCCTCAGTTGTTATTGATT 360
Do 301 ATCATTTGCGCCAAATTAGCTGTATAATTAGTCCTCAGTTGTTATTGATT 360
QY 361 GATAGTTGAGTTGAAGGTACAGAATTTC 392
Do 361 GATAGTTGAGTTGAAGGTACAGAATTTC 392

XX OS Candida albicans; strain SGV-1.
 XX WO200263397-A1.
 XX PD 11-MAY-2000.
 XX PT 01-NOV-1999; 99WO-NZ000179.
 XX PR 30-OCT-1998; 98CA-02249046.
 XX PR 30-OCT-1998; 98US-0106342P.
 XX PA (JANSEN PHARM NV.
 XX PI Luyten WHML, De Backer MD, Nelissen BJM, Poulter RTM;
 XX DR WPI; 2000-365640/31.
 XX PT Novel retrotransposon expression vectors useful for expressing an
 PT antigen, epitope or therapeutic agent, or detecting genes or the presence
 PT of Candida in a sample.
 XX PS Example 9; Fig 9; 204pp; English.
 XX CC The invention relates to novel retrotransposons from the yeast Candida
 CC albicans which have a copy number of 40-150, preferably 50-100 copies per
 genome. In particular, the invention relates to the novel C. albicans
 CC Ty1/copia retrotransposon pCal (AA57920), and to the integrated form of
 CC this retrotransposon, designated Tca2, and to the novel C. albicans
 CC retrotransposons 1-28. pCal was initially isolated from C. albicans
 CC h01042 and has a copy number of 50-100 copies per cell. It comprises
 CC identical 280 bp long terminal repeats (LTRs) and two open reading frames
 CC (ORFs). The first ORF encodes a gag (group antigen) protein, and the
 CC second ORF encodes a polyprotein (pol) consisting of an aspartate
 CC protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and
 CC pol ORFs of pCal are in the same reading frame, separated only by a
 CC termination codon (TGA). Translation of the pol ORF occurs through the
 CC occasional readthrough suppression of the stop codon, which is mediated
 CC by the formation of a pseudoknot within the gag-pol mRNA. The
 CC retrotransposons of the invention can be used as vectors for in vitro or
 CC in vivo transformation and expression. They can thus be used for the
 CC delivery and expression of a therapeutic, immunological or immunogenic
 CC molecule (e.g., an antigen) and may also be used for eliciting an
 CC immunological response in a host organism. They are therefore useful in
 CC genetic vaccine compositions and for gene therapy. Particularly where the
 CC use of retroviral vectors is unsafe or undesirable. Additionally, the
 CC retrotransposons may be used to generate transgenic animals, to detect
 CC the presence of Candida in a sample, to detect and disrupt genes, and to
 CC assign functions to nucleotide sequences. Sequences AA57923-AA57935
 CC represents TCA2 retrotransposon 5' regions from a variety of C. albicans
 CC strains. (Updated on 15-SEP-2003 to standardize OS field)
 XX SQ Sequence 392 BP; 136 A; 56 C; 72 G; 128 T; 0 U; 0 Other;
 XX Query Match 6.1%; Score 388.8; DB 3; Length 392;
 XX Best Local Similarity 99.5%; Pred. No. 7.9e-69;
 XX Matches 330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 XX PS Example 9; Fig 9; 204pp; English.
 XX CC The invention relates to novel retrotransposons from the yeast Candida
 CC albicans which have a copy number of 40-150, preferably 50-100 copies per
 genome. In particular, the invention relates to the novel C. albicans
 CC Ty1/copia retrotransposon pCal (AA57920), and to the integrated form of
 CC this retrotransposon, designated Tca2, and to the novel C. albicans
 CC h01042 and has a copy number of 50-100 copies per cell. It comprises
 CC identical 280 bp long terminal repeats (LTRs) and two open reading frames
 CC (ORFs). The first ORF encodes a gag (group antigen) protein, and the
 CC second ORF encodes a polyprotein (pol) consisting of an aspartate
 CC protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and
 CC pol ORFs of pCal are in the same reading frame, separated only by a
 CC termination codon (TGA). Translation of the pol ORF occurs through the
 CC occasional readthrough suppression of the stop codon, which is mediated
 CC by the formation of a pseudoknot within the gag-pol mRNA. The
 CC retrotransposons of the invention can be used as vectors for in vitro or
 CC in vivo transformation and expression. They can thus be used for the
 CC delivery and expression of a therapeutic, immunological or immunogenic
 CC molecule (e.g., an antigen) and may also be used for eliciting an

XX OS RESULT 6
 XX ID AAA57932 standard; DNA; 392 BP.
 XX AC AAA57932;
 XX DT 15-SEP-2003 (revised)
 XX DT 10-OCT-2000 (first entry)
 XX DE Candida albicans strain F16-2 Tca2 retrotransposon 5' region.
 XX KW Retrotransposon; pCal; Tca2; Ty1; copia; long terminal repeat; LTR;
 XX KW gag gene; group antigen; polyprotein; pol; aspartate protease; integrase;
 XX KW reverse transcriptase; RNaseH; pseudoknot; readthrough translation;
 XX KW stop codon suppression; gene delivery; gene therapy vector;
 XX KW genetic vaccine composition; immunogenic; transgenic animal; ds
 XX OS Candida albicans; strain F16-2.
 XX PN WO200263397-A1.
 XX PR 30-OCT-1998; 98CA-02249046.
 XX PR 30-OCT-1998; 98US-0106342P.
 XX PA (JANSEN PHARM NV.
 XX PI Luyten WHML, De Backer MD, Nelissen BJM, Poulter RTM;
 XX DR WPI; 2000-365640/31.
 XX PT Novel retrotransposon expression vectors useful for expressing an
 PT antigen, epitope or therapeutic agent, or detecting genes or the presence
 PT of Candida in a sample.
 XX PS Example 9; Fig 9; 204pp; English.
 XX CC The invention relates to novel retrotransposons from the yeast Candida
 CC albicans which have a copy number of 40-150, preferably 50-100 copies per
 genome. In particular, the invention relates to the novel C. albicans
 CC Ty1/copia retrotransposon pCal (AA57920), and to the integrated form of
 CC this retrotransposon, designated Tca2, and to the novel C. albicans
 CC h01042 and has a copy number of 50-100 copies per cell. It comprises
 CC identical 280 bp long terminal repeats (LTRs) and two open reading frames
 CC (ORFs). The first ORF encodes a gag (group antigen) protein, and the
 CC second ORF encodes a polyprotein (pol) consisting of an aspartate
 CC protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and
 CC pol ORFs of pCal are in the same reading frame, separated only by a
 CC termination codon (TGA). Translation of the pol ORF occurs through the
 CC occasional readthrough suppression of the stop codon, which is mediated
 CC by the formation of a pseudoknot within the gag-pol mRNA. The
 CC retrotransposons of the invention can be used as vectors for in vitro or
 CC in vivo transformation and expression. They can thus be used for the
 CC delivery and expression of a therapeutic, immunological or immunogenic
 CC molecule (e.g., an antigen) and may also be used for eliciting an

CC immunological response in a host organism. They are therefore useful in CC genetic vaccine compositions and for gene therapy, particularly where the CC use of retroviral vectors is unsafe or undesirable. Additionally, the CC retrotransposons may be used to generate transgenic animals, to detect CC the presence of *Candida* in a sample, to detect and disrupt genes, and to CC assign functions to nucleotide sequences. Sequences AAA57923-A57936 CC represents TCa2 retrotransposon 5', regions from a variety of *C. albicans* CC strains. (Updated on 15-SEP-2003 to standardise OS field)

XX SQ Sequence 392 BP; 136 A; 56 C; 72 G; 128 T; 0 U; 0 Other;

Query Match 6.1%; Score 388.8; DB 3; Length 392;

Best Local Similarity 99.5%; Pred. No. 7.9e-69;

Matches 390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGTGGTTGTCGACTATTGTCAGAGACTGTCAGAAATGATGGTTATTGA 60

Db 1 TGTGGTTGTCGACTATTGTCAGAGACTGTCAGAAATGATGGTTATTGA 60

QY 61 GATGGAAATTTCATCAGCACATCAGGATGACAGACTAACTATTTGTT 120

Db 61 GATGGAAATTTCATCAGCACATCAGGATGACAGACTAACTATTTGTT 120

QY 121 ATAATTAAGGTTGAAATACCAACATCCCGAAATATCAGAGATGAGGAGT 180

Db 121 ATAATTAAGGTTGAAATACCAACATCCCGAAATATCAGAGATGAGGAGT 180

QY 181 TCAATATAGGTTGAAATACCAACATCCCGAAATATCAGAGATGAGGAGT 240

Db 181 TCAATATAGGTTGAAATACCAACATCCCGAAATATCAGAGATGAGGAGT 240

QY 241 ACACGCTCAATCTCAGTAAGAAGTTATTCATCAGATTAGAGCTGATGAT 300

Db 241 ACACGCTCAATCTCAGTAAGAAGTTATTCATCAGATTAGAGCTGATGAT 300

QY 301 ATCATTTGTCGAAATTAGGTGTTATAATTAGTCCTCAGTTGTTATTGTT 360

Db 301 ATCATTTGTCGAAATTAGGTGTTATAATTAGTCCTCAGTTGTTATTGTT 360

QY 361 GATAGTTGGAGTTGAGGTACAGAATTTC 392

Db 361 GATAGTTGGAGTTGAGGTACAGAATTTC 392

RESULT 7

AAA57927 ID AAA57927 standard; DNA; 392 BP.

AC AAA57927;

DT 15-SEP-2003 (revised)

DT 10-OCT-2000 (first entry)

XX Candida albicans strain ATC-1 TCa2 retrotransposon 5' region.

XX Retrotransposon; pfal; Tca2; Ty1; copia; long terminal repeat; LTR;

KK gag gene; group antigen; polyprotein; pol; aspartate protease; integrase;

KK gag; gene; transcript; RNaseL; pseudoknot; readthrough translation;

KK reverse transcriptase; RNaseL; pseudoknot; readthrough translation;

KK stop codon suppression; gene delivery; gene therapy vector;

KK genetic vaccine composition; gene; immunogenic; transgenic animal; ds.

OS Candida albicans; strain ATC-1.

XX WO200263197-A1.

PD 11-MAY-2000.

XX 01-NOV-1999; 99WO-NZ000179.

XX 30-OCT-1998; 98CA-02249046.

PR XX (JANS) JANSSEN PHARM NV.

XX

XX Luyten WHML, De Backer MD, Nelissen BWM, Poulter RTM; PT WPI; 2000-365640/31.

XX Novel retrotransposon expression vectors useful for expressing an PT antigen, epitope or therapeutic agent, or detecting genes or the presence PT of *Candida* in a sample.

XX Example 9; Fig 9; 204pp; English.

XX The invention relates to novel retrotransposons from the yeast *Candida* albicans which have a copy number of 40-150, preferably 50-100 copies per CC genome. In particular, the invention relates to the novel *C. albicans* CC Tyl/copia retrotransposon TCa2, and to the integrated form of CC Tyl/copia retrotransposon TCa2, and to the novel *C. albicans* CC hgl012 and has a copy number of 50-100 copies per cell. It comprises CC identical 280 bp long terminal repeats (LTRs) and two open reading frames CC (ORFs). The first ORF encodes a gag (group antigen) protein, and the CC second ORF encodes a polypeptide (pol) consisting of an aspartate CC protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and CC pol ORFs of TCa2 are in the same reading frame, separated only by a CC termination codon (TGA). Translation of the pol ORF occurs through the CC occasional readthrough suppression of the stop codon, which is mediated CC by the formation of a pseudoknot within the gag-pol mRNA. The CC retrotransposons of the invention can be used as vectors for in vitro or CC in vivo transformation and expression. They can thus be used for the CC delivery and expression of a therapeutic, immunological or immunogenic CC molecule (e.g., an antigen) and may also be used for eliciting an CC immunological response in a host organism. They are therefore useful in CC genetic vaccine compositions and for gene therapy, particularly where the CC use of retroviral vectors is unsafe or undesirable. Additionally, the CC retrotransposons may be used to generate transgenic animals, to detect CC the presence of *Candida* in a sample, to detect and disrupt genes, and to CC assign functions to nucleotide sequences. Sequences AAA57923-A57936 CC represents TCa2 retrotransposon 5', regions from a variety of *C. albicans* CC strains. (Updated on 15-SEP-2003 to standardise OS field)

XX SQ Sequence 392 BP; 136 A; 56 C; 72 G; 128 T; 0 U; 0 Other;

Query Match 6.1%; Score 388.8; DB 3; Length 392;

Best Local Similarity 99.5%; Pred. No. 7.9e-69;

Matches 390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGTGGTTGTCGACTATTGTCAGAGACTGTCAGAAATGATGGTTATTGA 60

Db 1 TGTGGTTGTCGACTATTGTCAGAGACTGTCAGAAATGATGGTTATTGA 60

QY 61 GATGGAAATTTCATCAGCACATCAGGATGACAGACTAACTATTTGTT 120

Db 61 GATGGAAATTTCATCAGCACATCAGGATGACAGACTAACTATTTGTT 120

QY 121 ATAATTAAGGTTGAAATACCAACATCCCGAAATATCAGAGATGAGGAGT 180

Db 121 ATAATTAAGGTTGAAATACCAACATCCCGAAATATCAGAGATGAGGAGT 180

QY 181 TCAATATAGGTTGAAATACCAACATCCCGAAATATCAGAGATGAGGAGT 240

Db 181 TCAATATAGGTTGAAATACCAACATCCCGAAATATCAGAGATGAGGAGT 240

QY 241 ACACGCTCAATCTCAGTAAGAAGTTATTCATCAGATTAGAGCTGATGAT 300

Db 241 ACACGCTCAATCTCAGTAAGAAGTTATTCATCAGATTAGAGCTGATGAT 300

QY 301 ATCATTTGTCGAAATTAGGTGTTATAATTAGTCCTCAGTTGTTATTGTT 360

Db 301 ATCATTTGTCGAAATTAGGTGTTATAATTAGTCCTCAGTTGTTATTGTT 360

QY 361 GATAGTTGGAGTTGAGGTACAGAATTTC 392

Db 361 GATAGTTGGAGTTGAGGTACAGAATTTC 392

gag gene; group antigen; polyprotein; pol; aspartate protease; integrase; reverse transcriptase; RNaseH; Pseudoklenow; readthrough translation; stop codon suppression; gene delivery; gene therapy vector; genetic vaccine composition; immunogenic; transgenic animal; ds. Candida albicans; strain SA4-1. WO20026397-A1. 11-MAY-2000. 01-NOV-1999; 99WO-NZ000179. 30-OCT-1998; 98CA-02249046. 30-OCT-1998; 98US 0106342P. (JANCO) JANSSEN PHARM NV. Luyten WHML, De Backer MD, Nelissen BJM, Poulter RTM; WPI; 2000-36564/31. Novel retrotransposon expression vectors useful for expressing an antigen, epitope or therapeutic agent, or detecting genes or the presence of Candida in a sample. Example 9; Fig 9; 04pp; English.

QY	181	TCATATATATATCTTGCAATACTCTGCTTAATCTCGTTAATCTCACTATACAACTAGTCATAGT	241
Db	181	TCAATATATATCTTGCAATACTCTGCTTAATCTCACTATACAACTAGTCATAGT	241
Db	181	TCAATATATATCTTGCAATACTCTGCTTAATCTCACTATACAACTAGTCATAGT	241
Db	301	ATCATTTGTCGCAATTAGGTGTTGATAATTCAGTCTGAGATGTTATGATT	301
QY	301	ATCATTTGTCGCAATTAGGTGTTGATAATTCAGTCTGAGATGTTATGATT	301
Db	361	GATAGTTGAGGTTGAGGTCAGAAGATT	361
QY	361	GATAGTTGAGGTTGAGGTCAGAAGATT	361
Db	361	GATAGTTGAGGTTGAGGTCAGAAGATT	361
RESULT 11			
AAA57930			
ID	AAA57930	standard; DNA; 392 BP.	
XX			
AC	AAA57930;		
XX			
DT	15-SEP-2003	(revised)	
XX	10-OCT-2000	(first entry)	
DE	Candida albicans	strain SA4-2	Tca2 retrotransposon 5' region.
XX			
KW	Retrotransposon; pCal; Tca2; Ty1; copia; long terminal repeat; LTR; gag gene; group antigen; polyprotein; pol; aspartate protease; integrase; reverse transcriptase; RNaseH; pseudoknot; readthrough translation; stop codon suppression; gene delivery; gene therapy vector; genetic vaccine composition; immunogenic; transgenic animal; ds.		
KW			
KW			
OS	Candida albicans; strain	SA4-2.	
XX			
DN	WO200263397-A1.		
XX			
PD	11-MAY-2000.		
XX			
PF	01-NOV-1999;	99WO-NZ000179.	
XX			
PR	30-OCT-1998;	98CA-02249046.	
XX	30-OCT-1998;	98US-0106342P.	
PA	(JANCO)	JANSSEN PHARM NV.	
XX			
PI	Luyten WHML, De Backer MD, Nelissen BJM, Poulter RTM;		
XX			
DR	WPI; 2000-365640/31.		
PS			
XX			
CC	Novel retrotransposon expression vectors useful for expressing an antigen, epitope or therapeutic agent, or detecting genes or the presence of Candida in a sample.		
XX			
Example 9: Fig 9; 204pp; English.			
The invention relates to novel retrotransposons from the yeast <i>Candida albicans</i> which have a copy number of 40-150, preferably 50-100 copies per genome. In particular, the invention relates to the novel <i>C. albicans</i> Ty1/copia retrotransposon pCal (AAA57920), and to the integrated form of this retrotransposon, designated Tca2, and to the novel <i>C. albicans</i> retrotransposons 1-28. pCal was initially isolated from <i>C. albicans</i> b01102 and has a copy number of 50-100 copies per cell. It comprises identical 280 bp long terminal repeats (LTRs) and two open reading frames (ORFs). The first ORF encodes a gag (group antigen) protein, and the second ORF encodes a polyprotein (pol), consisting of an aspartate protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and pol ORFs of pCal are in the same reading frame, separated only by a termination codon (TGA). Translation of the pol ORF occurs through the occasional readthrough suppression of the stop codon, which is mediated by the formation of a pseudoknot within the gag-pol mRNA. The			

CC retrotransposons of the invention can be used for the delivery and expression of a therapeutic, immunological or immunogenic molecule (e.g., an antigen) and may also be useful for eliciting an immunological response in a host organism. They are therefore useful in genetic vaccine compositions and for gene therapy, particularly where the use of retroviral vectors is unsafe or undesirable. Additionally, the retrotransposons may be used to generate transgenic animals, to detect the presence of *Candida* in a sample, to detect and disrupt genes, and to assign functions to nucleotide sequences. Sequences *AA55923*-*AS57936* represents *TA2* retrotransposon 5' regions from a variety of *C. albicans* strains. (updated on 15-SEP-2003 to standardise OS field)

PR	30-OCT-1998;	98CA-0224946.
XX	30-OCT-1998;	98US-0106342P.
PA	(JANCO) JANSSEN PHARM NV.	
PT	Luyten WHML, De Backer MD, Nelissen BJM, Poulter RTM;	
XX	WPI; 2000-365640/31.	
XX	Novel retrotransposon expression vectors useful for expressing an antigen, epitope or therapeutic agent, or detecting genes or the presence of <i>Candida</i> in a sample.	
PS	Example 9; Fig 9; 204pp; English.	
XX	The invention relates to novel retrotransposons from the yeast <i>Candida albicans</i> which have a copy number of 40-150, preferably 50-100 copies per genome. In particular, the invention relates to the novel <i>C. albicans</i> Ty1/copia retrotransposon pCal (AA57920), and to the integrated form of this retrotransposon, designated TC2, and to the novel <i>C. albicans</i> retrotransposons 1-28. pCal was initially isolated from <i>C. albicans</i> h0142 and has a copy number of 50-100 copies per cell. It comprises identical 280 bp long terminal repeats (LTRs) and two open reading frames (ORFs). The first ORF encodes a gag (group antigen) protein, and the second ORF encodes a polypeptide (pol) consisting of an aspartate protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and pol ORFs of pCal are in the same reading frame, separated only by a termination codon (TGA). Translation of the pol ORF occurs through the occasional readthrough suppression of the stop codon, which is mediated by the formation of a pseudoknot within the gag-pol mRNA. The retrotransposons of the invention can be used as vectors for in vitro or in vivo transformation and expression. They can thus be used for the delivery and expression of a therapeutic, immunological or immunogenic molecule (e.g., an antigen) and may also be used for eliciting an immunological response in a host organism. They are therefore useful in genetic vaccine compositions and for gene therapy, particularly where the use of retroviral vectors is unsafe or undesirable. Additionally, the retrotransposons may be used to generate transgenic animals, to detect the presence of <i>Candida</i> in a sample, to detect and disrupt genes, and to assign functions to nucleotide sequences. Sequences AA57923-A57936 represents TC2a retrotransposon 5' regions from a variety of <i>C. albicans</i> strains. (Updated on 15-SEP-2003 to standardise OS field)	
XX	Sequence 392 BP; 138 A; 59 C; 71 G; 124 T; 0 U; 0 Other;	
SQ	Query Match 6.0%; Score 384; DB 3; Length 392; Best Local Similarity 98.7%; Pred. No. 7.4e-68; Matches 387; Conservative 0; Mismatches 5; Indels 0; Gaps 0; Matches 387; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
QY	1 TGTGCGTTGCACTTGTGCGAACTGATCATGAAATGAGTGTATTATGA 60	
DB	1 TGTGCGTTGCACTTGTGCGAACTGATCATGAAATGAGTGTATTATGA 60	
QY	61 GATGCGAAATTTCGCACTGAGTGGTGTGAGAACTAACTATGTTGAGT 120	
DB	61 GATGCGAAATTTCGCACTGAGTGGTGTGAGAACTAACTATGTTGAGT 120	
QY	121 ATATATAGGTTATGAAATCCACACCGAGATACACGGAGATAGAGGGAGTT 180	
DB	121 ATATATAGGTTATGAAATCCACACCGAGATACACGGAGATAGAGGGAGTT 180	
QY	241 ACGCGCTCAATCAGGAAAGATATTCATCAGATAGAGTCATAGT 300	
DB	241 ACGCGCTCAATCAGGAAAGATATTCATCAGATAGAGTCATAGT 300	
QY	301 ATCATATATCTGTGAAATAGGTTGATAATTCAGCTCGATGTTATTTATT 360	
DB	301 ATCATATATCTGTGAAATAGGTTGATAATTCAGCTCGATGTTATTTATT 360	
QY	301 AACATTTCGCCAAATAGGTTGATAATTCAGCTCGATGTTATTTATT 360	
DB	301 AACATTTCGCCAAATAGGTTGATAATTCAGCTCGATGTTATTTATT 360	
XX	RESULT 13	
XX	AA57935	
ID	AA57935 standard; DNA; 392 BP.	
XX	AC AA57935;	
XX	DE <i>Candida albicans</i> strain p30 TC2 retrotransposon 5' region.	
XX	XX Retrotansposon; pCal; TC2; Ty1; copia; long terminal repeat; LTR; gag gene; group antigen; polypeptide; pol; aspartate protease; integrase; reverse transcriptase; RNaseH; pseudoknot; readthrough translation; stop codon suppression; gene delivery; gene therapy vector; genetic vaccine composition; immunogenic; transgenic animal; ds.	
XX	OS <i>Candida albicans</i> ; strain p30.	
XX	XX WO200026397-A1.	
XX	XX PD 11-MAY-2000.	
XX	XX PF 01-NOV-1999; 99WO-NZ000179.	
XX	XX PR 30-OCT-1998; 98CA-0224946.	
XX	XX PR 30-OCT-1998; 98US-0106342P.	
XX	XX PA (JANCO) JANSSEN PHARM NV.	
XX	XX PT Luyten WHML, De Backer MD, Nelissen BJM, Poulter RTM;	
XX	XX DR WPI; 2000-365640/31.	
XX	XX PT Novel retrotransposon expression vectors useful for expressing an antigen, epitope or therapeutic agent, or detecting genes or the presence of <i>Candida</i> in a sample.	
XX	XX PT Novel retrotransposon expression vectors useful for expressing an antigen, epitope or therapeutic agent, or detecting genes or the presence of <i>Candida</i> in a sample.	
PS	Example 9; Fig 9; 204pp; English.	
XX	The invention relates to novel retrotransposons from the yeast <i>Candida albicans</i> which have a copy number of 40-150, preferably 50-100 copies per genome. In particular, the invention relates to the novel <i>C. albicans</i> Ty1/copia retrotransposon pCal (AA57920), and to the integrated form of this retrotransposon, designated TC2, and to the novel <i>C. albicans</i> retrotransposons 1-28. pCal was initially isolated from <i>C. albicans</i> h0142 and has a copy number of 50-100 copies per cell. It comprises identical 280 bp long terminal repeats (LTRs) and two open reading frames (ORFs). The first ORF encodes a gag (group antigen) protein, and the second ORF encodes a polypeptide (pol) consisting of an aspartate protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and pol ORFs of pCal are in the same reading frame, separated only by a termination codon (TGA). Translation of the pol ORF occurs through the occasional readthrough suppression of the stop codon, which is mediated by the formation of a pseudoknot within the gag-pol mRNA. The retrotransposons of the invention can be used as vectors for in vitro or in vivo transformation and expression. They can thus be used for the delivery and expression of a therapeutic, immunological or immunogenic molecule (e.g., an antigen) and may also be used for eliciting an immunological response in a host organism. They are therefore useful in genetic vaccine compositions and for gene therapy, particularly where the use of retroviral vectors is unsafe or undesirable. Additionally, the retrotransposons may be used to generate transgenic animals, to detect the presence of <i>Candida</i> in a sample, to detect and disrupt genes, and to assign functions to nucleotide sequences. Sequences AA57923-A57936 represents TC2a retrotransposon 5' regions from a variety of <i>C. albicans</i> strains. (Updated on 15-SEP-2003 to standardise OS field)	

XX
DE Candida albicans strain SC5-2 TCa2 retrotransposon 5' region.
XX
KW Retrotransposon; pCa2; Tyl; copia; long terminal repeat; LTR;
KW gag gene; group antigen; polyprotein; pol; aspartate protease; integrase;
KW reverse transcriptase; RNaseH; pseudoknot; readthrough translation;
KW stop codon suppression; gene delivery; gene therapy vector;
KW genetic vaccine composition; immunogenic; transgenic animal; ds.
XX
OS Candida albicans; strain SC5-2.
XX
PN WO20026397-A1.
XX
PD 11-MAY-2000.
XX
PP 01-NOV-1999; 99WO-N2000179.
XX
PR 30-OCT-1998; 98CA-02249046.
PR 30-OCT-1998; 98US-0106342P.
XX
PA (JANCO) JANSSEN PHARM NV.
XX
PI Luyten WHML, De Backer MD, Nelissen BJM, Poultier RTM;
XX
DR WPI; 2000-365640/31.
XX
PT Novel retrotransposon expression vectors useful for expressing an
PT antigen, epitope or therapeutic agent, or detecting genes or the presence
PT of Candida in a sample.
XX
PS Example 9, Fig 9; 204pp; English.
XX
CC The invention relates to novel retrotransposons from the Yeast Candida
CC albicans which have a copy number of 40-150, preferably 50-100 copies per
genome. In particular, the invention relates to the novel C. albicans
CC Ty1/copia retrotransposon pCa2 (AA57930), and to the integrated form of
CC this retrotransposon, designated TCa2, and to the novel C. albicans
CC retrotransposons 1-28. pCa2 was initially isolated from C. albicans
CC HOG142 and has a copy number of 50-100 copies per cell. It comprises
CC identical 280 bp long terminal repeats (LTRs) and two open reading frames
CC (ORFs). The first ORF encodes a gag (group antigen) protein, and the
CC second ORF encodes a polyprotein (pol) consisting of an aspartate
CC protease, integrase, reverse transcriptase (RT) and RNaseH. The gag and
CC pol ORFs of pCa2 are in the same reading frame, separated only by a
CC termination codon (TGA). Translation of the pol ORF occurs through the
CC occasional readthrough suppression of the stop codon, which is mediated
CC by the formation of a pseudoknot within the gag-pol mRNA. The
CC retrotransposons of the invention can be used as vectors for in vitro or
CC in vivo transformation and expression. They can thus be used for the
CC delivery and expression of a therapeutic, immunological or immunogenic
CC molecule (e.g., an antigen) and may also be used for eliciting an
CC immunological response in a host organism. They are therefore useful in
CC genetic vaccine compositions and for gene therapy, particularly where the
CC use of retroviral vectors is unsafe or undesirable. Additionally, the
CC retrotransposons may be used to generate transgenic animals, to detect
CC the presence of Candida in a sample, to detect and disrupt genes, and to
CC assign functions to nucleotide sequences. Sequences AAA57923-AA57935
CC represents TCa2 retrotransposon 5' regions from a variety of C. albicans
CC strains. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 392 BP; 137 A; 57 C; 71 G; 127 T; 0 U; 0 Other;

Query Match 5.9%; Score 380.8; DB 3;
Best Local Similarity 98.2%; Prid. No. 3.3e-67;
Matches 385; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TTTGGTTGTCGACTATTGTCGAGACTGATCAATGAAATGTTGTTATTGAA 60
Db 1 TGTGGTTGTCGACTATTGTCGAGACTGATCAATGAAATGTTGTTATTGAA 60
Qy 61 GATGGAATTTTTCATCAGCTAGTATGATCAGACTAACTTATGTTGAT 120
Db 61 GATGGAATTTTTCATCAGCTAGTATGATCAGACTAACTTATGTTGAT 120

Qy 121 ATTAATAGGGTATGAAATACACATCCCGAAATTCACAGAGTAGAGGGAGGTT 180
Db 121 ATAATATAGGGTATGAAATACACATCCCGAAATTCACAGAGTAGAGGGAGGTT 180
Qy 181 TCAATATATGTTGATATACTTCTGTTAACTTCTGTTAACTACAGACTAGGTT 240
Db 181 TCAATATATGTTGATATACTTCTGTTAACTACAGACTAGGTT 240
Qy 241 ACACGCTCAATCTCAGGTTAGAAAGTTTCTCATACACTGGCT 300
Db 241 ACACGCTCAATCTCAGGTTAGAAAGTTTCTCATACACTGGCT 300
Qy 301 AACATATTGTCGCAATTAGGTGTTAAATTGTTCTCGTCTCGAGTTATTGTT 360
Db 301 AACATATTGTCGCAATTAGGTGTTAAATTGTTCTCGTCTCGAGTTATTGTT 360
Qy 361 GATAGTTGAGTTGAGGTACAGAATTC 392
Db 361 GATAGTTGAGTTGAGGTACAGAATTC 392

Search completed: April 11, 2004, 14:02:32
Job time : 1504 secs

This Page Blank (uspto)

RESULT 4
 US-09-800-729-33
 ; Sequence 33, Application US/09800729
 ; GENERAL INFORMATION:
 ; APPLICANT: Ni et al. 32 Human secreted proteins
 ; TITLE OF INVENTION: 32 Human secreted proteins
 ; FILE REFERENCE: PZ044P1
 ; CURRENT APPLICATION NUMBER: US/09/800,729
 ; CURRENT FILING DATE: 2001-03-08
 ; PRIORITY APPLICATION NUMBER: PCT/US00/26013
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: 60/155,709
 ; PRIORITY FILING DATE: 1999-09-24
 ; NUMBER OF SEQ ID NOS: 217
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 33
 ; LENGTH: 2394
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-800-729-33

Query Match 0.9%; Score 57.2; DB 4; Length 2394;
 Best Local Similarity 52.0%; Pred. No. 0.0012; Gaps 0;
 Matches 128; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy 1133 GNAATGCTGATCATCTGATATGCTAGATGGTGCACATCCAGCACAAATGAA 1192
 Db 2148 GAACTCTTAATGATGCTTACTTTACTTTAAATTCAGAACATCCATAAAGGA 2207
 Qy 1193 CTTAACTCTTATGTTAGTTACCCAGGACCGAGAACAAAGGAAACAGGAGAC 1252
 Db 2208 AGGGGGTAACTTATGATGAGAAAGAAAGAAAGAAACTTCATGGATGA 2267
 Qy 1253 TCACTGGACACTCTAAAGAGAACCTGAAATAGAAAGAAATAGAACATCA 1312
 Db 2268 AAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAA 2327
 Qy 1313 AATCGATAGATAAGCTGAAAGAAAGAAAGAAAGAAACTTCATGGATGA 1372
 Db 2328 AAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAA 2387
 Qy 1373 AAAAC 1378
 Db 2388 AAAAAA 2393

RESULT 5
 US-10-204-708-52/C
 ; Sequence 52, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; FILE REFERENCE: 5013_1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; CURRENT FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06

RESULT 6
 US-09-014-969-14
 ; Sequence 14, Application US/09014969
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: Lavallie, Edward R.
 ; APPLICANT: Racine, Lisa A.
 ; APPLICANT: Werner, David
 ; APPLICANT: Tracey, Maurice
 ; APPLICANT: Spaulding, Vicki
 ; APPLICANT: Agostino, Michael J.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
 ; NUMBER OF SEQUENCES: 32
 ; NUMBER OF SEQUENCES: 32
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02140

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/014,969
 FILING DATE:

CLASSIFICATION: ATTORNEY/AGENT INFORMATION:
 NAME: Springer, Suzanne A.
 REGISTRATION NUMBER: 41, 323
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 876-5551
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2447 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-014-969-14

Query Match 0.9%; Score 56.4; DB 2; Length 2447;
 Best Local Similarity 53.8%; Pred. No. 0.0018;
 Matches 114; Conservative 1; Mismatches 97; Indels 0; Gaps 0;
 QY 1167 TTGACTTCCAGCACAAATATGACTTATCTTATGTTAGTTACCGCAGCAGA 1226
 Db 2203 TTGATTGTTGACTTAAATGACATGACAAATAACCTTGGAGAAMAAAARAAA 2262

QY 1227 AACCAAGGAAACCGAGGAGAACCTACATGGACAACTCAGAGAAGACCTGMAA 1286
 Db 2263 AAAAARAAAARAAAARAAAARAAAARAAAARAAAARAAAARAAAARAAA 2322

QY 1287 CAGAAGAGAAATAGAACATCCAAATCAGATAGCTAACGATAAGGTGAAAGAANG 1346
 Db 2323 AAAAGAAAAGACTTCACTGGATGAAACA 1378

QY 1347 AAAAGAAAAGACTTCACTGGATGAAACA 2414

Db 2383 AAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 2414

RESULT 7
 US-09-973-462-2
 ; Sequence 2, Application US/08973462B
 ; GENERAL INFORMATION:
 ; APPLICANT: DAUBRESES, PIERRE
 ; APPLICANT: DAUBRESES, PIERRE
 ; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
 ; FILE REFERENCE: 0660-0125-0 PCT
 ; CURRENT APPLICATION NUMBER: US/08/973, 462B
 ; CURRENT FILING DATE: 1998-02-06
 ; EARLIER APPLICATION NUMBER: PCT/FR96/00894
 ; EARLIER FILING DATE: 1996-06-12
 ; EARLIER FILING DATE: 1995-06-13
 ; NUMBER OF SEQ ID NOS: 29
 ; SEQ ID NO: 1
 ; LENGTH: 6152
 ; SOFTWARE: PatentIn Ver. 2.0
 ; LENGTH: 6152
 ; SEQ ID NO: 1
 ; TYPE: DNA
 ; ORGANISM: P. falciparum
 ; US-08-973-462-1

Query Match 0.9%; Score 56; DB 3; Length 6152;
 Best Local Similarity 45.9%; Pred. No. 0.0031; Mismatches 225; Indels 0; Gaps 0;
 Matches 191; Conservative 0; Mismatches 225; Indels 0; Gaps 0;
 QY 3869 ACTGTTGCTAAAGTTATGGAGATGAAATCTCTCTTATTTATCATTAGTGTATC 3928
 Db 533 AATGGTGAATTAGTCTAGTGTAGAGAATGAAAGAAATCTCTGACTTTAGAGAAGGA 592

QY 3929 ACTGAACTTGCTACAGCTGGAAATGATAGGAAATTACAGAACTAACGAACTCGCACITCACTG 3988
 Db 593 AATACATTAATGAACTTAAAGTGTAGTGTAGAAGAAGCAACAGAAATTTGACATTAAG 652

QY 3989 AAAATGAAATATCACCAGTGATAATGAAAGAAATATGAAATAATCCAAACACATT 4048
 Db 653 GAAATATTTATTAAGTAAATAGAAGAAGCAACAGAAATTTGACATTAAG 712

QY 4049 GAAAGTATCCTGCTGATAGAGATGGATGAAATTGAAACGATAATGTTGAAATT 4108
 Db 713 AATATATTTGGACAAATTCAGAAACAGAAAGTGTATCAGAAATGTACAGTAGT 772

QY 4109 GAGATGTTGAAATGACATGACATGACATGCTGAGCTAATCCACATCAGATGAAATT 4168
 Db 773 GATGAACTTTTGTAAATTAATGTTGAGTGTAAATGAGAGTAAAGAAAT 832

QY 4169 GATGTCAGATGATGAGAGITGATATAATCATAGAATGTCAGGAAAGAGAA 4228
 Db 833 ATTGAGGAAAGTCAGTAAATGACATTAATGAGTAAATGTTGAAATGTTCAA 892

QY 4708 TTACACTGTTATAGACTTGTGACTATAAGATTAGGAGATAATGGTGTGAATTAG 4767
 Db 367 TCCCCGGTGCAGAALATGACTACTATTCGACTCTTATTCGACTCTTACTTC 425
 Db 4758 AATGACAAATTCAACATTAGGCTGAGTAAATGCTCGCGTATCAAATGCGCT 4827
 QY 427 GTCGCAATTCTCACTTGATGGTTAAATGCCTCTGATGGAGACTTCAGAGA 486
 Db 4828 AATCCCAATTAGTCTTCCCTAAATCAGTACATTGAGAAACATGGGTT 4887
 QY 487 AGTTTATGTCGACTCTCTCTCTGATGTC 4944
 Db 4948 AAGAGTAACTGGAGACATGGTTTACTCAAGTTAACATGATGGTTACAT 5007
 QY 605 TATGTTGACTCTGCTCTGGATTCTAGACTCAAGTTAACATGATGGTTACAT 5067
 Db 5008 TCAATGAGTAAAGGGGACTAGTAAATATTAGTTATGATGATATCTTATGGT 5067
 QY 662 TATGTTGACTGAGTCGCGCTCGTATCTGCTTATGATGATGGTTACAT 721
 Db 5068 TGGAGTCACAAAAGTTGATATTGGGATCAATTGAGATCATTTGAGT 5127
 QY 722 TGTGATGACATGATGATGTTAGTTCAGCTTGAGAGAAAGTGGCTAGATGAA 781
 Db 5128 TAAGTGTGTTGAAATATCAATATCTGTTGATGAAATTGTAACGAACTGG 5187
 QY 782 GAAAGATTTGGGTATCTCAATATTCTCAAGAAATTACTTAAGGATTCAGAGG 841
 Db 5188 TATATTATTCAGAAATTCTCAAGAAATTACTTAAGGATTCAGACTAGATG 5247
 QY 842 TTAACCTTCTCACTGAAATTATGTCAGACATATCTGAGCAGACTAGTTGA 901
 Db
 RESULT 12
 US-09-254-776B-50
 ; Sequence 50, Application US/09254776B
 ; GENERAL INFORMATION:
 ; Patent No. 659359
 ; APPLICANT: Laten, Howard
 ; TITLE OF INVENTION: PLANT RETROVIRAL POLYNUCLEOTIDES AND METHODS FOR USE THEREOF
 ; CURRENT APPLICATION NUMBER: US/09/254,776B
 ; CURRENT FILING DATE: 1999-03-09
 ; NUMBER OF SEQ ID NOS: 86
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 50
 ; LENGTH: 2826
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: SIRE-1 genomic clone
 ; US-09-254-776B-50

Query Match 0.9%; Score 55.4; DB 4; Length 2826;
 Best Local Similarity 43.7%; Pred. No. 0.0032; Matches 343; Conservative 0; Mismatches 436; Indels 6; Gaps 2;
 Matches 343; Conservative 43.7%; Pred. No. 0.0032; Matches 448; Indels 6; Gaps 2;

QY 4484 AATGTCGGATGATAAGTCATGATGCTGACTAGAGAAATTAGATCAAAGAT 4543
 Db 1324 ACTGATGAGTCCTCGGATCAATGCTGAGAAGATGGAGATCAAGGATGAA 1383
 QY 4544 GTTTCGAGAAGTTCGAAATCCACCGGGTGGAGCTATATCTGGTGGTAC 4603
 Db 1384 GTTTCGAGCTGAGTCCTAGGCCGGAACTAAATGATGATGGACCAAGTGATC 1443
 QY 4604 ACTGAGAAATTGATCTCAAGGTTGGTGGAGATCACGTTGTTGTCATGCG 4663

QY 4724 CTTGAGTCACTCAGACTGTTACTCTGTTGATGCTCAATTCAGATCAAT 4783
 Db 1561 CTTGAGTCACTCAGACTGTTACTCTGTTGATGCTCAATTCAGCTGAC 1620
 QY 4784 TTAGCGTCAGTCGGCATCTAAATGCTCTTACTCTCATCAATCAATTAGTC 4843
 Db 1621 ATGGATGAGAGGCAATTGATGGATACCTGATGATGAAAGCTATGAGC 1880
 QY 4844 TTCTCTTAATCGTACCTTGAGAAACATTTGTTGTTATGAAACGCTCTGTC 4903
 Db 1681 CCAGGGATTTGATGATCCAACTCATCCAGATCATGATACAGCTCAGAGCTCTGC 1740
 QY 4904 TATGGTTAAACACGCGTTGATGTTACACTACATGATGTTACATCAAAAGCT 5023
 Db 1741 TATGATGAGCAAGCTGATGTTGATGAGCTTACAGAGTATGGAGAGC 5023
 QY 4964 ATGGTTTACTCAAGTTACACATGATGGTTTACACTGATGTTACATCAAAAGCT 5083
 Db 1801 CAAGGTTATGGAGGGGGATGACAGACCCCTT 5083
 QY 5024 TCAAGTATAATTAGTTATATGATGATTTCTATGTTGGAATTCAAAA 5083
 Db 1858 AATGATGATGAGCATAGATATCTGATGACATGCTGTTGGGGATGTCATGAG 1917
 QY 5084 GTTATGATATTGTTGGATCAATTGAGATCATTTGAGATCATTTGAGTA 5143
 Db 1918 ATGCTTGACATTTGTCACAGATGCAATTGAGATCATTTGAGTAAGGTTGAG 1977
 QY 5144 ATATCAATATCTGTTGATGAAATTGAAACCGATCTGGTTATTTATCTCAA 5203
 Db 1978 CTGATATTGTTGGATCAAGTGAAGCAGATGGAAGAATCCATATCCTTCACAA 2037
 QY 5204 GAAAATTCTCAAGAAATTACTTAGGATTCAACTAGATGACTCATGGGAAAC 5263
 Db 2038 AGCAAGTATGCAAGAACATTTGCAAGAAGTTGGATGGAAATGCCAGCCATAPAGA 2097
 QY 5264 ATACC 5268
 Db 2098 ACACC 2102

RESULT 13
 US-09-835-811-1
 ; Sequence 1, Application US/09835811
 ; GENERAL INFORMATION:
 ; Patent No. 6482936
 ; APPLICANT: HU, Song et al
 ; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, HUMAN SECRETED PROTEINS, AND NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND METHODS FOR USE THEREOF
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND METHODS FOR USE THEREOF
 ; FILE REFERENCE: C10_01228
 ; CURRENT APPLICATION NUMBER: US/09/835,811
 ; CURRENT FILING DATE: 2001-04-17
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1696
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-835-811-1

Query Match 0.9%; Score 55.2; DB 4; Length 1696;
 Best Local Similarity 57.6%; Pred. No. 0.0029; Matches 99; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Db 1508 TAGATTAAGAAAGAAAGAAACCTGAAATCAGAAAGAAATAGAAACATCCAAATCAGATATCGA 1567
 Qy 1267 TCAAAAAGAAGAACCTGAAATCAGAAAGAAATAGAAACATCCAAATCAGATATCGA 1326
 Db 1568 AAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAA 1627
 Qy 1327 TAAAGGTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1378
 Db 1628 AAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAA 1679

RESULT 14
 US-09-771-602D-1
 Sequence 1, Application US/08771602D
 GENERAL INFORMATION:
 PATENT NO. 5976795
 APPLICANT: Voytas, Daniel F.
 TITLE OF INVENTION: Retrotransposon and Methods
 NUMBER OF SEQUENCES: 51
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Greene, Winner and Sullivan, P.C.
 STREET: 5370 Manhattan Circle, Suite 201
 CITY: Boulder
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/771,602D
 FILING DATE: 20-DEC-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/010,869
 FILING DATE: 31-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Ferber, Donna M.
 REGISTRATION NUMBER: 33,878
 REFERENCE/DOCKET NUMBER: 8-96
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 499-8080
 TELEFAX: (303) 499-8089
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6660 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: not relevant
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: *Saccharomyces paradoxus*
 STRAIN: NRRL Y-17217
 FEATURE: CDS
 LOCATION: 1441..6321
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1228..6602
 OTHER INFORMATION: /function= "retrotransposon"
 OTHER INFORMATION: /product= "Ty5-6p"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 6352..6602
 OTHER INFORMATION: /function= "5' LTR of Ty5-6p"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 6352..6602

RESULT 15
 US-09-232-446B-1
 Sequence 1, Application US/09232446B
 PATENT NO. 6228647
 GENERAL INFORMATION:
 APPLICANT: Voytas, Daniel F.
 APPLICANT: Gai, Xiaowu
 TITLE OF INVENTION: Transposon Element Protein that Directs DNA
 FILE REFERENCE: 2-98
 CURRENT APPLICATION NUMBER: US/09/232,446B
 PRIOR APPLICATION NUMBER: US 60/071,383
 PRIOR FILING DATE: 1998-01-15
 NUMBER OF SEQ ID NOS: 26
 SEQ ID NO 1
 LENGTH: 6660
 TYPE: DNA
 ORGANISM: *Saccharomyces paradoxus*
 FEATURE: CDS
 LOCATION: (1441)..(6318)
 US-09-232-446B-1

Query Match Similarity 0.8%; Score 54; DB 3; Length 6660;
 Best Local Similarity 45.2%; Pred. No. 0.0067; Mismatches 289; Indels 3; Gaps 1;
 Matches 241; Conservative 0; Mismatches 289; Indels 3; Gaps 1;

Qy 4661 GGCACAGCAAAAGGAATATTGATTGACCTTTAGTGTAGTCACGTATA 4720
 Db 4981 GCCAAGGACATACCTCAAAAGCTGTTAGTACTATCAAGAACTTTCACCGTCATT 5040

Qy 4721 GATCTTGACTATAGATATTGACAATATAGTTGGAATTAGGATGACAATCAA 4780
 Db 5041 CGATAGCACTGTGATGATTCTGGCCTGCTAGTGCCCTGAACTAATAGTATA 5100

Qy 4781 CATTAGACGTGAGTCGGTATCTAAATGCCTTATTACTCATCATTCAATCCATT 4840
 Db 5101 CAGATGAGCTGACCGGTTCAAACTCAAAATGATGCGCTATAGTAAA 5160

Qy 4841 GCTCTTCCTCTAAATCAGTACCTTGAAAGAAAACATGTTGTTATGAAACGTCT 4900
 Db 5161 CAACCAACCGATTATTAAATGAAGTAATCCGACTATCTATGGAACTATACGGGGT 5220

Qy 4901 GTCTATGGTTAACAGTGGGTTGGATGGATCACACTTCAAAAGATGGCAA 4960
 Db 5221 ATGTAAGACTCAAGAACCCATTACTATGGACGACATATCACACTCTCA 5280

Qy 4961 GACATGGTTTACTCAAGTTTACAGTTTACAGTGTATTCACTATGAAAGAG 5020
 Db 5281 AAGATGGTTTCGTCGACATGAAGCGACATGGCTTAA-ActTGTCACAT 5337

Qy 5021 GGATCGTATAATATTAGTTTATGTTGATGATTTATGTTGAGCTTCAGAA 5080
 Db 5338 GATGGRCACATCTACATTGCTTACAGTACGAGACACTTCTGCTCCCTCG 5397

Qy 5081 AAAGTATGATAATTGGATGATGAGATTTGAGTTAGTAAAGTTTGT 5140
 Db 5398 AAATATAGACAGGTTAACAGAAACTAACAGAGTTACCTATGAGGAGCTGGT 5457

Qy 5141 GAAATCAATTATCTGGTTATGAAACCACTCTGTTAT 5193
 Db 5458 AAAGTACAAATCTGGCTTACATTCATTCAATGGACAT 5510

Search completed: April 11, 2004, 21:18:25
 Jcb time : 294 secs

Run on: April 11, 2004, 18:29:18 ; Search time 1445 Seconds
 (without alignments)
 16682.970 Million cell updates/sec

Scoring table: IDENTITY_NUC

Gappen 10.0 , Gapext 1.0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications NA: *

1: /cgn2_6/podata/2/pubpna/US07_PUBCOMB.seq: *
 2: /cgn2_6/podata/2/pubpna/PCT_NEW_PUB.seq: *
 3: /cgn2_6/podata/2/pubpna/US06_PUBCOMB.seq: *
 4: /cgn2_6/podata/2/pubpna/US06_PUBCOMB.seq: *
 5: /cgn2_6/podata/2/pubpna/US07_NEW_PUB.seq: *
 6: /cgn2_6/podata/2/pubpna/PCTUS_PUBCOMB.seq: *
 7: /cgn2_6/podata/2/pubpna/US08_NEW_PUB.seq: *
 8: /cgn2_6/podata/2/pubpna/PCT_NEW_PUBCOMB.seq: *
 9: /cgn2_6/podata/2/pubpna/US09_PUBCOMB.seq: *
 10: /cgn2_6/podata/2/pubpna/US09_PUBCOMB.seq: *
 11: /cgn2_6/podata/2/pubpna/US09_PUBCOMB.seq: *
 12: /cgn2_6/podata/2/pubpna/US09_NEW_PUB.seq: *
 13: /cgn2_6/podata/2/pubpna/US10_PUBCOMB.seq: *
 14: /cgn2_6/podata/2/pubpna/US10_PUBCOMB.seq: *
 15: /cgn2_6/podata/2/pubpna/US10_PUBCOMB.seq: *
 16: /cgn2_6/podata/2/pubpna/US10_NEW_PUB.seq: *
 17: /cgn2_6/podata/2/pubpna/US10_NEW_PUB.seq: *
 18: /cgn2_6/podata/2/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	92	1.4	1289	12	US-10-424-599-71662
2	89.8	1.4	5417	12	US-10-424-599-71662
3	84.2	1.3	4029	12	US-10-424-599-71662
4	84.2	1.3	4143	12	US-10-424-599-71662
5	83.8	1.3	13341	8	US-08-910-386A-1
6	79.6	1.3	78333	15	US-10-298-122-3
7	70.8	1.1	560	12	US-10-424-599-188696
8	70.6	1.1	2760	12	US-10-424-599-188338
9	68.4	1.1	3673778	14	US-10-312-841-1
10	68.2	1.1	7306	14	US-10-311-455-110
11	65.8	1.0	3073778	14	US-10-312-841-2
12	65.6	1.0	7628	12	US-10-424-599-188640
13	65.4	1.0	11729	14	US-10-311-455-858
14	65.2	1.0	681	9	US-09-22-830-856
15	65.2	1.0	12177	14	US-10-311-455-624

RESULT 1
 US-10-424-599-71662
 Sequence 71662, Application US/10424599
 Publication No. US20040031072A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa Thomas J
 APPLICANT: Kovacic David K
 APPLICANT: Zhou Yihua
 APPLICANT: Gao Yongwei
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 Title of Invention: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5323)B
 CURRENT APPLICATION NUMBER: US/10/424,599
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 285684
 SEQ ID NO 71662
 LENGTH: 1289
 TYPE: DNA
 ORGANISM: Glycine max
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (1)..(1289)
 OTHER INFORMATION: unsure at all n locations
 FEATURE:
 OTHER INFORMATION: unsure at all n locations
 OTHER INFORMATION: Clone ID: PAT_MRT3847_35724C.1
 US-10-424-599-71662
 Query Match 1.4%; Score 92; DB 12; Length 1289;
 Best Local Similarity 46.3%; Pred. No. 5.5e-08;
 Matches 376; Conservative 0; Mismatches 430; Indels 6; Gaps 2;
 Sequence 1, Appli
 Sequence 2, Appli
 Sequence 3, Appli
 Sequence 11, Appli
 Sequence 12, Appli
 Sequence 13, Appli
 Sequence 14, Appli
 Sequence 15, Appli
 Sequence 16, Appli
 Sequence 17, Appli
 Sequence 18, Appli
 Sequence 19, Appli
 Sequence 20, Appli
 Sequence 21, Appli
 Sequence 22, Appli
 Sequence 23, Appli
 Sequence 24, Appli
 Sequence 25, Appli
 Sequence 26, Appli
 Sequence 27, Appli
 Sequence 28, Appli
 Sequence 29, Appli
 Sequence 30, Appli
 Sequence 31, Appli
 Sequence 32, Appli
 Sequence 33, Appli
 Sequence 34, Appli
 Sequence 35, Appli
 Sequence 36, Appli
 Sequence 37, Appli
 Sequence 38, Appli
 Sequence 39, Appli
 Sequence 40, Appli
 Sequence 41, Appli
 Sequence 42, Appli
 Sequence 43, Appli
 Sequence 44, Appli
 Sequence 45, Appli
 Sequence 46, Appli
 Sequence 47, Appli
 Sequence 48, Appli
 Sequence 49, Appli
 Sequence 50, Appli
 Sequence 51, Appli
 Sequence 52, Appli
 Sequence 53, Appli
 Sequence 54, Appli
 Sequence 55, Appli
 Sequence 56, Appli
 Sequence 57, Appli
 Sequence 58, Appli
 Sequence 59, Appli
 Sequence 60, Appli
 Sequence 61, Appli
 Sequence 62, Appli
 Sequence 63, Appli
 Sequence 64, Appli
 Sequence 65, Appli
 Sequence 66, Appli
 Sequence 67, Appli
 Sequence 68, Appli
 Sequence 69, Appli
 Sequence 70, Appli
 Sequence 71, Appli
 Sequence 72, Appli
 Sequence 73, Appli
 Sequence 74, Appli
 Sequence 75, Appli
 Sequence 76, Appli
 Sequence 77, Appli
 Sequence 78, Appli
 Sequence 79, Appli
 Sequence 80, Appli
 Sequence 81, Appli
 Sequence 82, Appli
 Sequence 83, Appli
 Sequence 84, Appli
 Sequence 85, Appli
 Sequence 86, Appli
 Sequence 87, Appli
 Sequence 88, Appli
 Sequence 89, Appli
 Sequence 90, Appli
 Sequence 91, Appli
 Sequence 92, Appli
 Sequence 93, Appli
 Sequence 94, Appli
 Sequence 95, Appli
 Sequence 96, Appli
 Sequence 97, Appli
 Sequence 98, Appli
 Sequence 99, Appli
 Sequence 100, Appli
 Sequence 101, Appli
 Sequence 102, Appli
 Sequence 103, Appli
 Sequence 104, Appli
 Sequence 105, Appli
 Sequence 106, Appli
 Sequence 107, Appli
 Sequence 108, Appli
 Sequence 109, Appli
 Sequence 110, Appli
 Sequence 111, Appli
 Sequence 112, Appli
 Sequence 113, Appli
 Sequence 114, Appli
 Sequence 115, Appli
 Sequence 116, Appli
 Sequence 117, Appli
 Sequence 118, Appli
 Sequence 119, Appli
 Sequence 120, Appli
 Sequence 121, Appli
 Sequence 122, Appli
 Sequence 123, Appli
 Sequence 124, Appli
 Sequence 125, Appli
 Sequence 126, Appli
 Sequence 127, Appli
 Sequence 128, Appli
 Sequence 129, Appli
 Sequence 130, Appli
 Sequence 131, Appli
 Sequence 132, Appli
 Sequence 133, Appli
 Sequence 134, Appli
 Sequence 135, Appli
 Sequence 136, Appli
 Sequence 137, Appli
 Sequence 138, Appli
 Sequence 139, Appli
 Sequence 140, Appli
 Sequence 141, Appli
 Sequence 142, Appli
 Sequence 143, Appli
 Sequence 144, Appli
 Sequence 145, Appli
 Sequence 146, Appli
 Sequence 147, Appli
 Sequence 148, Appli
 Sequence 149, Appli
 Sequence 150, Appli
 Sequence 151, Appli
 Sequence 152, Appli
 Sequence 153, Appli
 Sequence 154, Appli
 Sequence 155, Appli
 Sequence 156, Appli
 Sequence 157, Appli
 Sequence 158, Appli
 Sequence 159, Appli
 Sequence 160, Appli
 Sequence 161, Appli
 Sequence 162, Appli
 Sequence 163, Appli
 Sequence 164, Appli
 Sequence 165, Appli
 Sequence 166, Appli
 Sequence 167, Appli
 Sequence 168, Appli
 Sequence 169, Appli
 Sequence 170, Appli
 Sequence 171, Appli
 Sequence 172, Appli
 Sequence 173, Appli
 Sequence 174, Appli
 Sequence 175, Appli
 Sequence 176, Appli
 Sequence 177, Appli
 Sequence 178, Appli
 Sequence 179, Appli
 Sequence 180, Appli
 Sequence 181, Appli
 Sequence 182, Appli
 Sequence 183, Appli
 Sequence 184, Appli
 Sequence 185, Appli
 Sequence 186, Appli
 Sequence 187, Appli
 Sequence 188, Appli
 Sequence 189, Appli
 Sequence 190, Appli
 Sequence 191, Appli
 Sequence 192, Appli
 Sequence 193, Appli
 Sequence 194, Appli
 Sequence 195, Appli
 Sequence 196, Appli
 Sequence 197, Appli
 Sequence 198, Appli
 Sequence 199, Appli
 Sequence 200, Appli
 Sequence 201, Appli
 Sequence 202, Appli
 Sequence 203, Appli
 Sequence 204, Appli
 Sequence 205, Appli
 Sequence 206, Appli
 Sequence 207, Appli
 Sequence 208, Appli
 Sequence 209, Appli
 Sequence 210, Appli
 Sequence 211, Appli
 Sequence 212, Appli
 Sequence 213, Appli
 Sequence 214, Appli
 Sequence 215, Appli
 Sequence 216, Appli
 Sequence 217, Appli
 Sequence 218, Appli
 Sequence 219, Appli
 Sequence 220, Appli
 Sequence 221, Appli
 Sequence 222, Appli
 Sequence 223, Appli
 Sequence 224, Appli
 Sequence 225, Appli
 Sequence 226, Appli
 Sequence 227, Appli
 Sequence 228, Appli
 Sequence 229, Appli
 Sequence 230, Appli
 Sequence 231, Appli
 Sequence 232, Appli
 Sequence 233, Appli
 Sequence 234, Appli
 Sequence 235, Appli
 Sequence 236, Appli
 Sequence 237, Appli
 Sequence 238, Appli
 Sequence 239, Appli
 Sequence 240, Appli
 Sequence 241, Appli
 Sequence 242, Appli
 Sequence 243, Appli
 Sequence 244, Appli
 Sequence 245, Appli
 Sequence 246, Appli
 Sequence 247, Appli
 Sequence 248, Appli
 Sequence 249, Appli
 Sequence 250, Appli
 Sequence 251, Appli
 Sequence 252, Appli
 Sequence 253, Appli
 Sequence 254, Appli
 Sequence 255, Appli
 Sequence 256, Appli
 Sequence 257, Appli
 Sequence 258, Appli
 Sequence 259, Appli
 Sequence 260, Appli
 Sequence 261, Appli
 Sequence 262, Appli
 Sequence 263, Appli
 Sequence 264, Appli
 Sequence 265, Appli
 Sequence 266, Appli
 Sequence 267, Appli
 Sequence 268, Appli
 Sequence 269, Appli
 Sequence 270, Appli
 Sequence 271, Appli
 Sequence 272, Appli
 Sequence 273, Appli
 Sequence 274, Appli
 Sequence 275, Appli
 Sequence 276, Appli
 Sequence 277, Appli
 Sequence 278, Appli
 Sequence 279, Appli
 Sequence 280, Appli
 Sequence 281, Appli
 Sequence 282, Appli
 Sequence 283, Appli
 Sequence 284, Appli
 Sequence 285, Appli
 Sequence 286, Appli
 Sequence 287, Appli
 Sequence 288, Appli
 Sequence 289, Appli
 Sequence 290, Appli
 Sequence 291, Appli
 Sequence 292, Appli
 Sequence 293, Appli
 Sequence 294, Appli
 Sequence 295, Appli
 Sequence 296, Appli
 Sequence 297, Appli
 Sequence 298, Appli
 Sequence 299, Appli
 Sequence 300, Appli
 Sequence 301, Appli
 Sequence 302, Appli
 Sequence 303, Appli
 Sequence 304, Appli
 Sequence 305, Appli
 Sequence 306, Appli
 Sequence 307, Appli
 Sequence 308, Appli
 Sequence 309, Appli
 Sequence 310, Appli
 Sequence 311, Appli
 Sequence 312, Appli
 Sequence 313, Appli
 Sequence 314, Appli
 Sequence 315, Appli
 Sequence 316, Appli
 Sequence 317, Appli
 Sequence 318, Appli
 Sequence 319, Appli
 Sequence 320, Appli
 Sequence 321, Appli
 Sequence 322, Appli
 Sequence 323, Appli
 Sequence 324, Appli
 Sequence 325, Appli
 Sequence 326, Appli
 Sequence 327, Appli
 Sequence 328, Appli
 Sequence 329, Appli
 Sequence 330, Appli
 Sequence 331, Appli
 Sequence 332, Appli
 Sequence 333, Appli
 Sequence 334, Appli
 Sequence 335, Appli
 Sequence 336, Appli
 Sequence 337, Appli
 Sequence 338, Appli
 Sequence 339, Appli
 Sequence 340, Appli
 Sequence 341, Appli
 Sequence 342, Appli
 Sequence 343, Appli
 Sequence 344, Appli
 Sequence 345, Appli
 Sequence 346, Appli
 Sequence 347, Appli
 Sequence 348, Appli
 Sequence 349, Appli
 Sequence 350, Appli
 Sequence 351, Appli
 Sequence 352, Appli
 Sequence 353, Appli
 Sequence 354, Appli
 Sequence 355, Appli
 Sequence 356, Appli
 Sequence 357, Appli
 Sequence 358, Appli
 Sequence 359, Appli
 Sequence 360, Appli
 Sequence 361, Appli
 Sequence 362, Appli
 Sequence 363, Appli
 Sequence 364, Appli
 Sequence 365, Appli
 Sequence 366, Appli
 Sequence 367, Appli
 Sequence 368, Appli
 Sequence 369, Appli
 Sequence 370, Appli
 Sequence 371, Appli
 Sequence 372, Appli
 Sequence 373, Appli
 Sequence 374, Appli
 Sequence 375, Appli
 Sequence 376, Appli
 Sequence 377, Appli
 Sequence 378, Appli
 Sequence 379, Appli
 Sequence 380, Appli
 Sequence 381, Appli
 Sequence 382, Appli
 Sequence 383, Appli
 Sequence 384, Appli
 Sequence 385, Appli
 Sequence 386, Appli
 Sequence 387, Appli
 Sequence 388, Appli
 Sequence 389, Appli
 Sequence 390, Appli
 Sequence 391, Appli
 Sequence 392, Appli
 Sequence 393, Appli
 Sequence 394, Appli
 Sequence 395, Appli
 Sequence 396, Appli
 Sequence 397, Appli
 Sequence 398, Appli
 Sequence 399, Appli
 Sequence 400, Appli
 Sequence 401, Appli
 Sequence 402, Appli
 Sequence 403, Appli
 Sequence 404, Appli
 Sequence 405, Appli
 Sequence 406, Appli
 Sequence 407, Appli
 Sequence 408, Appli
 Sequence 409, Appli
 Sequence 410, Appli
 Sequence 411, Appli
 Sequence 412, Appli
 Sequence 413, Appli
 Sequence 414, Appli
 Sequence 415, Appli
 Sequence 416, Appli

Query Match 1-4%; Score 89.8; DB 12; Length 5417; Best Local Similarity 46.8%; Pred. No. 3e-07; 0; Mismatches 322; Indels 0; Gaps 0;

QY 4606 TGAGAAATTGATCTCTAAGGGTGTGGAAATCACGTGCTGCATGCCA 4665
 Db 417 AACCAGCTCACTTCTGATGCTTAAACAACTAACAGCAAGCTGTGTA 473

QY 4666 CAGACAAAGGAATTGGATTGACCCCTTGTGTAGTCACCTGTATAGCT 4725
 Db 474 AGGATGCTAACGTTGGTGTGATTTGAGACTTTCGCTTGTGAGCT 533

QY 4726 TGTGACTATAGATTGACATATAAGTGTGATTAAGACATT 4785
 Db 534 AGAACCTAACCAACGGCGCTCTGTCACAAAGGTGGATPATACATCA 593

QY 4786 AGACGTCGAGTCGGCTATAAATGGCTTATPACTCAATCCATTAGCT 4845
 Db 594 GGATGTAATCAGCCTTCTTAATGATGACTTGAGGAGAAATTGTTGAGCAAC 653

QY 4846 TCCTCTTAATCAGTACCTTGAACAAACCATGTTGTTATGAAAGCTGTC 4905
 Db 654 TGAAGGGTCTGTCAAGACAGAGGAAGTATGATTAAGAAAGCTGA 713

QY 4906 TGGTAAACAGCTGGTTGGAAATCACATCAAAAGAGTATGGAGAT 4965
 Db 714 TGGTAAACAGGACCGCGCTATGGATAGCAGAATGTCACACTGGGAAC 773

QY 4966 TGGTTTACTCAAGTTCACAGTGGTTATTCACATGATGTTGAGGGTC 5025
 Db 774 AGGCTCAAAAGTAACTTAAAGTGTGAACTTATATGAGAGACTATGTA 833

QY 5026 AGTAATPATATTAGGTATATGATGATCTTATGGTGGAGTCAAAAGT 5085
 Db 834 GCTGTA-GTATCCTATGTTGAGGAGACGGGGCA 890

QY 5086 TATGTTAAACAGTGGGTTGGATGATCCTCAAGAGTATGGAGAC 5145
 Db 891 TATGAGACAGTCAAAGAGATGAGATGCTTGTGAACTGAGACTGGAGGT 950

QY 5146 ATCAATTATCTGGTATTAATTCTGTAACCAATCTGGTATTTCTCAAGA 5205
 Db 951 GACATCTTCGGTATGGGATCAAACAGAATATTCTTCCAGGA 1010

QY 5206 AAATTCTCGAAATTCTAAGGATCAACTAGTAGTGAATCTGGAAACAT 5265
 Db 1011 AAAATCTGAAATGAAATTCTCAAGAATCTGGATGGTCAACCAATGCA 1070

QY 5266 ACCCTGGATTCCGAATGCAAAATGAAAG 5297
 Db 1071 TOCAATGATCAAAGGGAGTTTACAAAG 1102

RESULT 2
 US-10-424-599-103057
 ; Sequence 103057, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovacic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; APPLICANT: Zhou Yihua
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 10357
 ; LENGTH: 5417
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE: OTHER INFORMATION: Clone ID: PAT_MRP13847_64078C.1
 ; OTHER INFORMATION: US-10-424-599-103057

Query Match 1-4%; Score 89.8; DB 12; Length 5417; Best Local Similarity 46.8%; Pred. No. 3e-07; 0; Mismatches 322; Indels 0; Gaps 0;

QY 4664 AACAGCAAAAGGAATTGGATTGACCTTGTGTTAGTTACCTGTATAGT 4723
 Db 3760 AAAGGTATAGCAGACATATGGAGTTGATGTTAGTCAAGTGTGTTGACCGGTGCCCC 3819

QY 4724 CTTGTCACATAAGATATGACATAATAGGTGTGAAATTGAGAACTCACAT 4783
 Db 3820 ATGGAAACCATCGTCTCTTCTGCGCTCAATGAAGTGGAAATTTCAG 3879

QY 4784 TTAGAGTCGAGTCGGTCAATGCTTATGCTTATACATCATTCAATCAATTGTC 4843
 Db 3880 CTGATGTAATCCGATTCATAATGGCTATCTGAAGAAATGCTATGIGCA 3939

QY 4844 TTTCCTTAATCAGTACCTTGAAGAAACCATGTTGTTATGAAACGTTGTC 4903
 Db 3940 CCAATGGTTTGTCTGTCAGGAGTCAAAGAAAGTGTGAATGACAAGGGTTC 3999

QY 4904 TATGGTTAAACAGTGGGTTGGATGATCCTCAAGAGTATGGAGAC 4963
 Db 4900 TATGGCTAAAGCAGCAGCAGGCGCATGCAACTCTGCAATGACAAGTTCAGAC 4059

QY 4964 ATGGTTTACTCAAGTTCACAGTGGTTATTCACATGATGTTTATTCACATGAAATGAGGG 5023
 Db 4960 ATGGGTGTTGTCGGTCAATGATGTCCTTATGTTAACTTATATGTT 4119

QY 5024 TCACTTATGTTATGTTATGTTATGTTATGTTATGTTGAGTTCAGA 5083
 Db 4120 GATGCTTATTATGTTCTTATGTTGAGTACCTTATCTTACCGCAATAACCAAT 4179

QY 5084 GTTATGATAATTCTGGTCAATGAGATCATTTGAGTTAAGTGTGGTGA 5143
 Db 4180 TTTGTAAGGCTCAGGCTCAGGTCCTGTCGTGAAATGATGACAGATGGACTC 4239

QY 5144 ATATCAAATTCTGGTATGAACTTCGAACAAAGCAATCTGTTATTCACAA 5203
 Db 4240 ATGTCATATTCTGGGATGGAAGTGAAGCAACGGAGATGTTCTTGTCA 4229

QY 5204 GAAATTCTCGAAATTCTAGGATCAACTAGTAGTGAATCTGGAAAC 5263
 Db 4300 GAAAGTACACAAAGAGTGTGAGAAATTATGTTGATGTCATCCGGTGAAC 4339

QY 5264 ATACC 5268
 Db 4360 ACACC 4364

RESULT 3
 US-10-424-599-42489
 ; Sequence 42489, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovacic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; APPLICANT: Zhou Yihua
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 42489
 ; LENGTH: 4029
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE: OTHER INFORMATION: Clone ID: PAT_MRP13847_1383669C.1
 ; OTHER INFORMATION: US-10-424-599-42489
 ; OTHER INFORMATION: US-10-424-599-103057

Best Local Similarity	46.3%	Pred. No.	3.3e-06	Matches	352	Conservative	0	Mismatches	403	Indels	6	Gaps	2																																																																																										
Qy	4496	TCGATTAAGTCATGATGTCGAACTAGAAATTAGATCAAAGATGTTAGAAGAA	4555	Db	2521	TGGCTAGAGAAGCTATGAAAGGAGACTCGATCCATAGAGAACTGACTTGAAGAA	2580	TYPE: DNA	ORGANISM: Glycine max	FEATURE: OTHER INFORMATION: Clone ID: PAT_MRT3847_138370C.1	US-10-424-599-42491	Query Match	1.3%	Score 84.2; DB 12; Length 4143;																																																																																									
Qy	4556	GTCCTCATTCACCGGTGAGGACTATCTATGGGTGGAACATCTGAGAAATT	4615	Db	2581	GTGATTACAAAGGAAGAAAGCAATGGTGAATGAAAGTAAAGTAAAGT	2640	Best Local Similarity	46.3%	Pred. No.	3.4e-06	Matches	352	Conservative	0	Mismatches	403	Indels	6	Gaps	2																																																																																		
Qy	4616	GATTCTCAAGGCTGTTGGAAATCGGTGCTGTCATGGACAAAG	4675	Db	2641	ATCC---AAGGTCAGTGTGAGTACAGAACAGTAAAGGAAATT	2697	Qy	4676	GAATAATTGGTTAGACCTTTAGTGTAGTCACCGTTAGATTTGACTATA	4735	Db	2698	CAAGACATGGTGTGATTAGGAGTACAGTACAGAACAGTAAAGGAAATT	2757	Qy	4736	AGATTATGACAAATAAGGTGAGTGGAAATGACAACTCACATTAAGCTCGAG	4795	Db	2758	AGAGTGTGGAGCTCATGCAAGTGCAGTGCAGGCGAACTCAACTTGTAA	2817	Qy	4796	TCGGGTATCTAACTCCTATTACTCATTCATCAATTGCTTCCCTAA	4855	Db	2818	TCTGCCTCTRAATGGAAATGGAGAGGAGGTAGTGTGAAACCGAGGTT	2877	Qy	4856	TCGACTCTTGAGAAACATGTTGTTATGAACTGAGAATGAGCTGTGCTATGGTTAA	4915	Db	2878	GTGAAACAGGAGTGAAGAAAGTACTGAATGAGAAAGCTGTGTTGAA	2937	Qy	4916	AGTCGGTTGGAACTACATCAAGAGTATGGAGACATTGTTACT	4975	Db	2938	CAAGTCACAGCTGTTGGAAAGAATGAGTAACTACAGGCTATG	2997	Qy	4976	CAAGTTTACATGATGGTTATTCACTGAAATGAGGATCAGIATAT	5035	Db	2998	AGATGTTCTGACATGGATGTTGAACTGAGTCAGGGGGTGTATGATA---	3054	Qy	5036	TAGGTATATGTTGATGATATTCTATGGTGGAACTTCAAAAGTTATGAT	5095	Db	3055	CTATGCTCTATGTTGATGATTTGTTGTTACTGGAACTATGAACTGAA	3114	Qy	5096	TGGTGGATCATTTGAGAGACATTGAGTTAAAGTGTGTTGAAATATCAATAT	5155	Db	3115	TAAAGGAAATGAGTGGACTTTGAAATGACAGACCTAGGAGATCTACTT	3174	Qy	5156	CTGGTATGATTGTAACCGATCTGGTATATTATCTAAGAAATTCTC	5215	Db	3175	CTAGGATTGATTCAAATACTAGTGTGTTGCTACTGTACAGAAATGTT	3234	Qy	5216	AAGAATTACTAAGATTCAAATAGTGTACCTGATAG	5256	Db	3235	TCTGATGTTACTCGAGAGATCAATGCAACACTGTAATGG	3275	RESULT 4	US-10-424-599-42491	Sequence 42491, Application US/10424599	Publication No. US20040031072A1	GENERAL INFORMATION:	APPLICANT: La Rosa Thomas J	APPLICANT: Kovacic David K	APPLICANT: Zhou Yihua	APPLICANT: Cao Yongwei	TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With	FILE REFERENCE: 38-21 (53223)B	CURRENT APPLICATION NUMBER: US/10-424,599	CURRENT FILING DATE: 2003-04-28	NUMBER OF SEQ ID NOS: 28564	SEQ ID NO 42491	LENGTH: 4143
Qy	5155	CTGGTATGATTGTAACCGATCTGGTATATTATCTAAGAAATTCTC	5215	Db	3239	AGAAATTACTAAGATTCAAATAGTGTACCTGATAG	3288	Qy	5216	AGAAATTACTAAGATTCAAATAGTGTACCTGATAG	5256	Db	3349	TCTGATGTTACTCGAGAGATCAATGCAACACTGTAATGG	3389	RESULT 5	US-08-10-166A-1	Sequence 1, Application US/08910386A	Publication No. US20020092041A1	GENERAL INFORMATION:	APPLICANT: Ronald, Pamela C.	APPLICANT: Wang, Guo-Liang	APPLICANT: Song, Wen-Yueng	APPLICANT: Hubert, Scot	APPLICANT: Richter, Todd	APPLICANT: Cao Yongwei	FILE REFERENCE: 38-21 (53223)B	CURRENT APPLICATION NUMBER: US/10-424,599	CURRENT FILING DATE: 2003-04-28	NUMBER OF SEQ ID NOS: 28564	SEQ ID NO 42491	LENGTH: 4143																																																																							

TITLE OF INVENTION: Procedures and Materials for Conferring
 TITLE OF INVENTION: Disease Resistance in Plants
 NUMBER OF SEQUENCES: 53
 CORRESPONDENCE ADDRESS:
 ADDRESS: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/910,386A
 FILING DATE: 13-AUG-1997
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Bastian, Kevin L.
 REGISTRATION NUMBER: 34,774
 REFERENCE/DOCKET NUMBER: 023070-058950US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 13341 base pairs
 STRAIN: IR8521
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT: 11
 MAP POSITION: 11Q, RG103
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 2367..4205
 OTHER INFORMATION: /product= "receptor kinase-like protein"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 4201..5071
 OTHER INFORMATION: /note= "retrofit", a copia-like,
 OTHER INFORMATION: transposon-like element"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 4484..8821
 OTHER INFORMATION: /product= "retrofit"
 OTHER INFORMATION: /gene= "gag/pol"
 FEATURE:
 NAME/KEY: intron
 LOCATION: 9915..1712
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 10020..10975
 OTHER INFORMATION: /note= "Krispie", transposon-like
 OTHER INFORMATION: element"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 12626..12750
 OTHER INFORMATION: /note= "Pop-012", transposon-like
 OTHER INFORMATION: element"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 13040..13248
 OTHER INFORMATION: /note= "Ds-rice2", transposon-like
 OTHER INFORMATION: element"
 US-08-910-386A-1

Query Match 1.3%; Score 83.9; DB 8; Length 13341;
 Best Local Similarity 46.4%; Pred. No. 7.1e-06;保守性 0; Mismatches 345; Conservative Matches 345; Matched 345; Indels 6; Gaps 2;
 Matches 345; Mismatches 0; Indels 6; Gaps 2;
 QY 4496 TGGATTAAGTAAATGATGCTGAATCTAGAAATTAGTCAAAGATCTTACGAGA 4555
 Db 7328 TGGAGAGTGCATGGAACTGAGTAACTGATAAATGACATGGCACTA 7387
 QY 4556 GRTCCCAATCCACCGGTGAGGCTATCTATGGGTGGGACACTGTGAGA 4615
 Db 7388 GRTCCCATGAGAAAGGACAATATCATGGGTAAATGGGTATAGATTAAG 7447
 QY 4676 GAAAATGGATTACCCCTTATGGTTAGTTAGTGTAGTCACCTGTATAGATTTGACTATA 4735
 Db 4616 GATTCTCAAGCTTGTGGAAATCAGTGTGTTGTCATGGCTACAGACAAAG 4675
 QY 7448 AAGGAGATGGACACTGTATGATAGATCACAGTAGCTGTGAGTGTG 7504
 Db 7448 AGATATTCTGTCCTTGCTGCTCTAGAGGTGGAGCTTAGAGTTAGATTCAG 7624
 QY 7505 CAAGGATATGGTATCATTAGAAGTACTCATCAACATTAGCTGCAACAACTCCGGTT 7564
 Db 74736 AGATTGTACATAATAGGTGIGATGGATTAGAATGACAACTTACGCTGAG 4795
 Db 7565 AGATATTCTGTCCTTGCTGCTCTAGAGGTGGAGCTTAGAGTTAGATTCAG 7624
 QY 47796 TCGGCGATCATTAACCTCTATTACTCACTCAACATTAGCTGCAACAACTCCGGTT 7624
 Db 7225 AATGCCTCTCTCATGGCTTACGTTAGAGAAGCTTACGCTGAG 4795
 QY 4856 TCACTCCTTGAGAAGAACATGGTTGTTATGGAACTGTTGTGTTAA 4915
 Db 7745 CAGCACCAAGGGCGGTGATTCCAGGCTAGTAAAGAACCTGTGTTAGCTGTT 7801
 QY 4976 CAGTTTACAAATGGTTTATTCATGATAAGGATAAGGGGAGCTAGTATAT 5035
 Db 7802 GAGCTCAAGGCTTACCTCTATTCTTAAAGGGAGGTATGGCTAA 7744
 QY 5036 TTAGGTTATGGTATGATATTCTTATGGTGTGGAGTCACAAAGATTGATAAT 5035
 Db 7862 GTTTGTTATGGTATGTTATGTTATGGTGTGGCTACAGAGGCAACTACAGCA 7921
 QY 5096 TTTGTTATGTTATGGAGAGCATTTGAGTTAAGTGTGTTGGAAATATCAATT 5155
 Db 7922 CTCTGAGGACTTAACAAAGGCTGACTTAAAGGTTGGGACCTGCACTTC 7981
 QY 5156 CTGGATTGATTGTTGAAACCGATCGTTATTTATCAGAAATTCTC 5215
 Db 7982 CTGGATTGAGGTAACATAAAGTTCAATGGGTATCTGACTCAAGAGAAGTGC 8041
 QY 5216 AAGAATTACTTAAAGGATTCAA 5238
 Db 8042 ATGACTCTGCTAAGGAGTAA 8064

RESULT 6
 US-10-298-122-3/C
 ; Sequence 3, Application US/10298122
 ; Publication No. US2003022124A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gmitter, Frederick G
 ; APPLICANT: Deng, Zhanbo
 ; APPLICANT: Zhang, Hongbin
 ; TITLE OF INVENTION: CITRUS TRISTIZA VIRUS AND METHODS OF USE
 ; FILE REFERENCE: 5053-220
 ; CURRENT APPLICATION NUMBER: US/10/298,122
 ; CURRENT FILING DATE: 2002-11-15
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 3

RESULT 7
 US-10-424-599-118696
 Sequence 118696, Application US/10424599
 Publication No. US2004031072A1
 GENERAL INFORMATION:

LENGTH: 78333
 TYPE: DNA
 ORGANISM: artificial
 FEATURE:
 OTHER INFORMATION: *Poncirus trifoliata* and *citrus grandis* hybrid
 US-10-298-122-3

Query Match 1.2%; Score 79.6; DB 15; Length 78333;
 Best Local Similarity 46.1%; Pred. No. 0. 0.0011; Matches 385; Conservative 0; Mismatches 439; Indels 12; Gaps 0; Matches 385; Conservative 0; Mismatches 439; Indels 12; Gaps 0

QY 4482 ATATATATCCTCCGGATGGATAAAGTCATGAATGCTGACTAGAGAATTAATGCAAG 45942 ATGTTTACGAGAAGTCATGCCAACGGGTGAGCCTATATCTAGGGTTGGTAC 4644529 AATATGATGAATCTTGGATTATGCCATGCAAGAGAGGAGTTGAACTTGAAG 464469 ATGTTGCGGAGTTAGTCCTAATCCGGATATAATCCATTATAGTCATAATGGTAT 464470 4602 ATACTGAGAAATTGATCTCTCAAAGGTGTCCTGG--AAATCACGTGTGTGTC 4644809 TTAGAACAAATGGATGAATCCGGTGTGTTAATAAATAGCTAGATAGTCCTC 4644959 ATGCCAACAGACAAAGAAAATTGGATTATGACCCTTAGTGTAGTTACCTGTA 4744949 AAGGTACGACCAAGAGAGAGGATGATTGATGAAACCTTG--CACCTGAG 474719 TAGATCTGTGACTATAAGATATTGAAATAATAGGTGTGATTAAGGAATGACAA 4744295 CAGATTTGAACTCATAGGATGTTGAGCATATGAGTCATAAGGTTTATTGTC 474779 AACATTAGACGCGAGTGGCGCTATCTAAATGCCCTATTACTCTCAATCCATT 4844235 ATCAAATGGATGTCAAAGTGCCTTCTTAATGGTACATTATGGAGGTGTTGTA 484939 ATGCTTTCTCTTAATCAGTACCTTGAAGAAACCATGTTGGTATGAAACCTT 4844175 AACACACCTCTGCTTGAATGAAATTTCAGACATGTTGAGTATTCAGAAG 484899 CTGTCATGGTTAACAGTCGGTTGGAATGGTATCACATCAAAAGGTATGG 4944115 CTTGATGGTTAACAGACACCTAGAGTTGGTATGATGATGCTTAATTTGTTG 494959 AAGACATGGTTTACTCAAGTTTACAACTATGGTTATTCTACATGTTGAG 505019 AGGGATCTGTAATATATTAGGTATATGTCATGATGATTCTTGTGAAAGTTCAC 5044055 TGGTAAATGATTTCT--AATGGAAAGGGACACRACTCTTTGTTAAGCATANGA 4343998 ATCAAGACATACTTGTCAATATAGTGTGATGATTATTGTTACTATG 435079 AAAAGTATGTTGATTTGIGGATCATGAGATCATTTGAGTTAAGTGTG 5143938 AGTTGTTGTAAGATTTCACTATGATGAGTTGAGTTGAGTGTGAAAGTGTG 515139 GTCGAATATCAAAATTCTGGATATGATTGTTAAACGGATCTGGTTATTATT 51439378 GAGAGTCAAGGATTCTCTGGACTCAATCAAACAAAGGGAGGAATTTCATAA 435199 CTCGAAGAAATTCTCAGAAATTCTAGGATTCAAACTAGATGACTCTATGGG 5243818 ACCAGCCAGTGTGAGAGATCTACTCAAAGATGTTGTTGATATGACGGCA 435259 AAAACACACCCCTGATGTCGAAATGACAATATGAAAGTGTGCAATTTCGTA 53443758 AAGCACTCTATGAGGACCACTCAAGCTGGACAAAGCTGAGAAGGTAAGGAA 43703

APPLICANT: La Rosa Thomas J
APPLICANT: Kovacic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424, 599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 118696
LENGTH: 560
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_78191C.1
US-10-424-599-118696

```

Query Match          1.1%; Score 70.8; DB 12; Length 560;
Best Local Similarity 51.7%; Pred. No. 0.0057; 0; Mismatches
Matches 187; Conservative 187; Indels 3; Gaps 1;
QY 4868 AAGAAAACCATGTGSTATGAAACGTCGTATGGTAAACAGTCGGTTC 4927
QY 45 AACCTTAATCAGTCCTTAACTGAAAGGTATATGGTAAACAGCCCTAGG 104
QY 4928 GAAATGTTACACTATAAAGAGGATGGAGACATGTTTACTCAAGTTTACAC 4987
QY 105 GCTTGTATGACTTGTAGTAATCTTGTAGCAAGGGTTCAAAAGTTAAG 161
QY 4988 AATGATGGTTATTCAGTGTATAGAGATGAGATGATACTCTTACTACAGTAT 5047
D9 162 GTTGTACTAACCTTTTATTAAAGAATGAGATGATACTCTTACTACAGTAT 221
D9 5048 GTTGTATGATTTTGTGTCACATATGATCTCTTGCACAAAAATTTCAGAT 5107
QY 222 GTTGTATGATTTTGTGTCACATATGATCTCTTGCACAAAAATTTCAGAT 281
QY 5108 TTGAGAGATCATTTGAGTAAAGTGTGGAATATCAATATCTGTGTTGAA 5157
QY 282 ATGCAAAATGAAATTGAAATGTCATGATGCGTGTAACTTTCTTGACTCAA 341
D9 5168 TTTGTTAAACCGTATCTGGTTATATTCAGAAATTCTCAGAAATCTCAGAA 5227
D9 342 ATGAAAGCAACAAAGAATGGTATTATCAGTCATCAAATATGCAAAAGACGTT 401
QY 5228 AA 5229
QY 402 CA 403
D9 402 CA 403

RESULT 8
US-10-424-5399-108338/c
; Sequence 108338 ; Application US/10424599
; Publication No. US20040311072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: SOY Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 3B-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 108338
; LENGTH: 2760
; TYPE: DNA
; FEATURE: ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_68844C.1
US-10-424-5399-108338

```


RESULT 11
 US-10-312-841-2/c
 ; Sequence 2, Application US/10312841
 ; Publication No. US20030186277A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Epigenomics AG
 ; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
 ; FILE REFERENCE: E01/1208/WO
 ; CURRENT APPLICATION NUMBER: US/10/312,841
 ; CURRENT FILING DATE: 2002-12-30
 ; NUMBER OF SEQ ID NOS: 2
 ; SEQ ID NO 2
 ; LENGTH: 3673778
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (379615)
 ; US-10-312-841-2

Query Match 1.0%; Score 65.8; DB 14; Length 3673778;
 Best Local Similarity 52.3%; Pred. No. 0.39; Gaps 0;
 Matches 145; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

Qy 1170 ACATTCAGGACAAATAGAACTTAATCTTATGTTAGTTACCGACACAGAGAAC 1229
 Db 3287077 AGCAACTTAATATAAACCAACCTTAATCTCAAAAAA 3287018

Qy 1230 CAAAGGAAACCCAGGGAGACTCACTGGACACTCTCAAGAGAACCTGAA 1289
 Db 3287017 AAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAA 3286958

Qy 1290 GAAAGGAATAGAAACATCCTAACTCAGTAAGATAAGGAAAGAAAGAA 1349
 Db 3286957 AAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAA 3286898

Qy 1350 AGAAAACACTCACTGGATGAAACGGTGGCTCTATTAATGTTAGATA 1409
 Db 3286897 AAAAAGAAAAAAAGAAAAAAAGAAAAAAATAACTTCTCATCAATCAAA 3286838

Qy 1410 TACATAATTGGCAACCGCGTTCGCTAGAAA 1446
 Db 3286837 CATATTAATCAATTAACCTTAAACCAATAAAAGAAAAA 3286801

RESULT 12
 US-10-424-599-108640
 ; Sequence 108640, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

Query Match 1.0%; Score 65.6; DB 12; Length 7628;
 Best Local Similarity 45.1%; Pred. No. 0.021; Gaps 0;
 Matches 371; Conservative 0; Mismatches 439; Indels 12; Gaps 3;

Qy 4484 ATGTATCGGGATGGATAAAGTCATGATGCTGAATCTAGAGAAATTAGTC 4543
 Db 4263 ACTGATGATGTTCTGATCACTCTATGCAAAAGAAATTGGGCAATTC 4322

Qy 4544 GTTTACGAGAGGTCCATTCACCGCTGTTGAMGCCTATPATCTATGGG 4603
 Db 4323 GTCGGGGCTAGTTCTAGACCAAGGAACTAATGTTAGCTGACATGCTC 4382

Qy 4604 ATGGAAATGATGCTCTCAAGGTGTTGTCG--GAATCAGTGTGTCAT 4660
 Db 4383 AAGACAAACAAATGAGGAAGGTGTTAACCGAAACAAAGCCAGACTTC 4442

Qy 4661 GGAACAGACAAAGGAAATTGGATTACCTTTATGTTAGTGTACCTTA 4720
 Db 4443 GGTACACTCAGATGGGGTAGCTTGA-----TGAATCTTCGCCCTGT 4495

Qy 4721 GATCTTGAGCTATAAGTATGACATAATGGTTGATTTGAGACAACTC 4780
 Db 4497 AGACTTGAGTCATCAGCATACTTGTTAGCTGTTACCTCTCAATCAGTGT 4556

Qy 4781 CATTAACTGGAGGGATCTAAAGGCCTPATATCTCATCAATCTCAATT 4840
 Db 4557 CAGATGGATGTAAGAGGGCGCTCTGATGATGAAAGGTCTATGGAG 4616

Qy 4841 GCTTTCCTAAATGAGTCACTTGAAGAAACCATGTTGCTTGAAGAAAGT 4900
 Db 4617 CAGCCAAAGGGATTCGAGTCACATCTACGATCATGATACAGCTCAAGAGCT 4676

Qy 4901 GCTATGGTTAAACGTCGGGTTGGATGTTACACTATCAAAGATGTTGA 4960
 Db 4677 CTCATGAGTGAAGGAGCTCAACAGAGCTTGTGTTGAGGTCAAAAGT 4736

Qy 4961 GACATGGTTTACTGAGTTACACATGAGTTACATGGTTACATGGATATGAG 5020
 Db 4737 CGCAGGGTATAGGAGGGAAATGACAGACTCTCTGTCAAACAGATGCTGA 4795

Qy 5021 GGTCAGTTATGTTATGTTATGTTATGATGATCTTATGTTGAGGTCACAA 5080
 Db 4797 AACTTGAGATA--GCACAGATATGTTGAGCATTTGAGTTAAGTGTGT 4853

Qy 5081 AAAGTTGATGTTGAGATGTTGAGCATTTGAGAGCATTTGAGTTAAGTGTGT 5140
 Db 4854 GAGATGCTTCGACATGTTGTCACAGATGAACTGTTGAGGATGAGTC 4913

Qy 5141 GAATTCATATCTGGATGATGTTGAAACGAACTGGTATTTCT 5200
 Db 4914 GAGCTGATTTCTGGACTTCAGTGAGCAGATGGAGGACTCCATTCCTCA 4973

Qy 5201 CAAAGAAATTCAGAAATTCTAGGTTGAGCTCAACTAGATGACTCATGGAA 5260
 Db 4974 CAAGCAGGTATGCAAGAACATGTCAGAGTGGATGGAGATGCCAGCATAAA 5033

RESULT 13
US-10-311-455-868/c
; Sequence 868, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determination of Cytosine Methylation
; FILE REFERENCE: 5013_1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIORITY FILING DATE: 2001-07-02
; PRIORITY APPLICATION NUMBER: DE 10032529.7
; PRIORITY FILING DATE: 2000-06-30
; PRIORITY APPLICATION NUMBER: DE 10043826.1
; PRIORITY FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 631
; SEQ ID NO: 26
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-868

Query Match 1.0%; Score 65.4; DB 14; Length 11729;
Best Local Similarity 53.3%; Pred. No. 0.029; Mismatches 121; Indels 0; Gaps 0;
Matches 138; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
Qy 1171 CATTCCAGCACAAATATGACTTATCTATGTTACGACAGAGAACCC 1230
Db 2878 CAACTTAAACACACACAAACGAAACCCATCTAAAGAAATAAACGACCA 2819
Qy 1231 AAAGGAAACGAGAGAGAACTCAATCTCAAAGAGAACCTGAAATCAG 1290
Db 2818 AAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAA 2759
Qy 1291 AAAGGAAATAAGAACATCCAAACAGATAACGTAAGGTGAAAGAAAGAAA 1350
Db 2758 AAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAA 2699
Qy 1351 AGAAAACCTCACTGGATGAAACAGGTGCTGTTCTTAATGTGTAATGAT 1410
Db 2658 AAAAAGAAAAAAAGAAAAAAAGAAAAAAACTCTACACCCAAACTTAATTCCAAAC 2639
Qy 1411 ACTATAATGCGAGCAACC 1429
Db 2638 CCAAACTCTATTAACCC 2620

RESULT 14
US-09-822-830A-26/c
; Sequence 26, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fecteau, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS

Query Match 1.0%; Score 65.2; DB 14; Length 12177;
Best Local Similarity 50.6%; Pred. No. 0.032; Mismatches 153; Indels 0; Gaps 0;
Matches 157; Conservative 0; Mismatches 153; Indels 0; Gaps 0;
Qy 1201 TATGTGTTATCCAGCACCGAGAACCAAGGAAACGAGGAGACTCTGGA 1260
Db 3594 TAATTAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAA 3535
Qy 1261 ACATCTCAAGAGAGACCTGAAATCAAGAGAAATAGAAATCTCAATCAGA 1320
Db 3534 AACAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAA 3475

Query Match 1.0%; Score 65.2; DB 9; Length 681;
Best Local Similarity 58.1%; Pred. No. 0.008; Mismatches 83; Indels 0; Gaps 0;
Matches 115; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
Qy 1188 ATGAACTAACTATGTTGTTACACACCAAGAGAACCCAAAGGAGAG 1247
Db 456 ATGAACTAACTATGTTGTTACACACCAAGAGAACCCAAAGGAGAG 397
Qy 1248 AGAACTACTGGACATCTCAAAGAGAACCTGAAATCAGAGAGAAATGAAAC 1307
Db 396 AAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAA 337
Qy 1308 ATCCAAATCAGATAACGATAAGGTGAAAGAAAGAAAGAAAGAAACTTCACTGG 1367
Db 336 AAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAA 277
Qy 1368 ATGAAAGAACGGTGTG 1385
Db 276 AAAAAGAAAAAGGGCTG 259

RESULT 15
US-10-311-455-624/c
; Sequence 624, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determination of Cytosine Methylation
; FILE REFERENCE: 5013_1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIORITY FILING DATE: 2001-07-02
; PRIORITY APPLICATION NUMBER: PCT/EP01/07537
; PRIORITY FILING DATE: 2000-06-30
; PRIORITY APPLICATION NUMBER: DE 10032529.7
; PRIORITY FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO: 624
; LENGTH: 12177
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-624

Query Match 1.0%; Score 65.2; DB 14; Length 12177;
Best Local Similarity 50.6%; Pred. No. 0.032; Mismatches 153; Indels 0; Gaps 0;
Matches 157; Conservative 0; Mismatches 153; Indels 0; Gaps 0;
Qy 1201 TATGTGTTATCCAGCACCGAGAACCAAGGAAACGAGGAGACTCTGGA 1260
Db 3594 TAATTAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAA 3535
Qy 1261 ACATCTCAAGAGAGACCTGAAATCAAGAGAAATAGAAATCTCAATCAGA 1320
Db 3534 AACAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAA 3475

Search completed: April 12, 2004, 01:18:39
Job time : 1460 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Om nucleic - nucleic search, using sw model

Run on: April 11, 2004, 13:02:12 ; Search time 9925 seconds
(without alignments)
19531.234 Million cell updates/sec

Title: US-09-430-590E-3
Perfect score: 6426
Sequence: 1 tggttgttgtgactattt.....agaagaatttatccatca 6426

Scoring table: IDENTITY_NUC
Gappen 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
listing first 45 summaries

Database : EST:*

1: em_estba:*

2: em_estham:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_ntc:*

9: gb_esti:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estcm:*

16: em_eston:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

ALIGNMENTS

RESULT 1
CNS07BJD
LOCUS CNS07BJD 1017 bp DNA linear
DEFINITION T7 end of clone BC0A012E08 of library BC0A from strain CBS 767 of
Debaryomyces hansenii, genomic survey sequence.
ACCESSION AL437951
VERSION GI:12221354
KEYWORDS GSS.
SOURCE Debaryomyces hansenii (anamorph: Candida famata)
ORGANISM Dikarya; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Debaryomyces.
REFERENCE 1 (bases 1 to 1017)
AUTHORS Soudet,J.J., Aigle,M., Artigauenave,F., Blandin,G.,
Boletin-Fukunara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,C., Durans,P., Lescigne,A., Llorente,B.,
Malpertuy,A., Neuveglise,C., Olier-Kalogeropoulos,O., Potier,S.,
Saurin,N., Tekai,F., Toffano-Noguel,C., Wesołowski-Louvel,M.,
Wincker,P. and Wisselbach,J.

Genomic exploration of the hemiascomycetous yeasts: 1. A set of
Yeast species for molecular evolution studies
Le singe,A., Casaregola,S., Neuveglise,C., Bon,E., Nguyen,H.,
Artigauenave,F., Wincker,P. and Gaillardin,C.

Result No. Score % Query Length DB ID Description
1 174.6 2.7 1017 29 CNS07BJD AL437951 T7 end of
2 142.4 2.2 905 29 CNS7ACK AL436555 T3 end of
3 127.2 2.0 838 29 CNS7ARP BI920632 EST540579
4 103.6 1.6 781 12 BI920632

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

TITLE	Genomic exploration of the hemiascomycetous yeasts: 14.
JOURNAL	Debaromyces hansenii var. hansenii FEMS Lett. 487 (1), 82-86 (2000)
MEDLINE	20584724
PUBMED	11152889
REFERENCE	3 (bases 1 to 1017)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail : secr@genoscope.cnrs.fr - Web : www.genoscope.cnrs.fr)
COMMENT	This GSS is part of a random genomic sequencing program of thirteen yeast species: <i>Saccharomyces bayanus</i> var. <i>uvvarum</i> , <i>Saccharomyces exigua</i> , <i>Saccharomyces servazzii</i> , <i>Zygosaccharomyces rouxii</i> , <i>Saccharomyces kluveri</i> , <i>Kluyveromyces thermotolerans</i> , <i>Kluyveromyces lactis</i> var. <i>lactis</i> , <i>Kluyveromyces marxianus</i> var. <i>marxianus</i> , <i>Pichia angusta</i> , <i>Debaromyces hansenii</i> var. <i>hansenii</i> , <i>Pichia sorbophila</i> , <i>Candida tropicalis</i> and <i>Yarrowia lipolytica</i> . Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See 5 kb were prepared and both extremities were sequenced. See the other extremity of this sequence and for the sequence of the other extremity of this insert.
FEATURES	Location/Qualifiers
source	1. . 1017
	/organism="Debaromyces hansenii"
	/mol-type="genomic DNA"
	/strain="CBS 767"
	/variety="hansenii"
	/db_xref="txxon:4959"
	/clone="BC0AA12008"
	/clone_id="BC0AA"
	/note="end : T7"
	<1..>1017
	/note="Part of putative transposable element"
	/evidence=not_experimental
ORIGIN	
Query Match	2.7%; Score 174.6; DB 29; Length 1017;
Best Local Similarity	53.2%; Pred. No. 5.3e-21;
Matches	412; Conservative 1; Mismatches 380; Indels 2; Gaps 2;
QY	CAAGAGGTATTACATGCCGTCACAAATGCACTTAATGATCCGGATGATAA 4503
Db	CACTCGGAAATTAACTATGTTATCGAAAGTGTGATTACAGATAACCTGGAGCA 303
QY	4504 GTCAATGAACTGACTGAGAAATTAGATCAAAGTGTGTTACGAGAACTTCAT 4563
Db	304 GTCAATGAACTGAGGATTAGATACATTAGAAATATGAGTATACCGTGGAA 363
QY	4564 TCCACCGGIGTGAACCCATATCATGGTGGTACACTGAGAATGATCTC 4623
Db	364 TCCTAAAATGCAACCTTCACACTTACCTTGTTCTACACATAATTAGACCT 423
QY	4624 CAAGGTTGTTGGAAATCACGGTGTGTGTCATGGAAACAGACAGAAGAAATT 4683
Db	424 CAAGAAGTTGAGTAAACACAGTCGGTGTGACGGCTTGGAAATGCAATGA 483
QY	4684 GGATTGACCTTGTAGTTAGTGTACCGTGTATAGACTTGCTTGACTATAAGATTT 4743
Db	484 ACACITGATACCTCGAGGTGTCTCTCTGTTGATGATTACATGATTCTGTTACT 543
QY	4744 GACAATPATAGTTGAGAATGGATGAGCAATTCACTTAAACGGTGGCTTA 4803
Db	544 TACAGGATACAGT-TGAATGATGGGCGATACATCTCTGTTATCTCCGATA 602
QY	4804 TCTAAATGCCCTATTAATCTCAATCAATTATGTTCTCTCTAAATCAGTAC 4863
Db	603 TTACATGCGATATGCGATATGAGAAATCATATTGTTAACCACTCTGGATCGA 662
QY	4864 TTGAAACAAACCATGTTGTTATGGAACTGCTGTCTGTTATGGTAAAGTCGG 4923
Db	663 TATTGATTCTGTAATGTTGCAAACTCTGTTATGGAAACGACAGCAGTGGTT 722
QY	4924 TTGGATGGTATCACATCAAAGAGTATGGAGAGACATGGTTACTCAAGTT 4983
RESULTS	2
LOCUS	CNS07ACX
DEFINITION	905 bp DNA linear GSS 08-JUL-2001 T3 end of clone BC0AA003C05 of library BC0AA from strain CBS 767 of Debaromyces hansenii, genomic survey sequence.
ACCESSION	AI436423
VERSION	AI36423.1
GSS	GI:12219836
KEYWORDS	
SOURCE	
ORGANISM	<i>Debaromyces hansenii</i> (anamorph: <i>Candida famata</i>)
SACCHAROMYCETES	<i>Eukaryota</i> ; <i>Fungi</i> ; <i>Ascomycota</i> ; <i>Saccharomycotina</i> ; <i>Saccharomycetes</i> ; <i>Saccharomycetales</i> ; <i>Saccharomycetaceae</i> ; <i>Debaromyces</i> .
REFERENCE	1 (bases 1 to 905)
AUTHORS	Souillet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Boilini,Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durress,P., Lepingle,A., Illorente,B., Malpertuy,A., Neuveglise,C., Olier,Kalogeropoulos,O., Potier,S., Saurin,W., Tekla,F., Toffano-Noguie,C., Wescowolski-Louvel,M., Wincker,P. and Weissbach,J.
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
JOURNAL	FEMS Lett. 487 (1), 3-12 (2000)
MEDLINE	20584711
PUBMED	11152876
REFERENCE	2 (bases 1 to 905)
AUTHORS	Lepingle,A., Casaregola,S., Neuveglise,C., Bon,E., Nguyen,H., Artiguenave,F., Wincker,P. and Gaillardin,C.
TITLE	Genomic exploration of the hemiascomycetous yeasts: 14.
JOURNAL	Debaromyces hansenii var. hansenii
MEDLINE	FEMS Lett. 487 (1), 82-86 (2000)
PUBMED	20584724
REFERENCE	3 (bases 1 to 905)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail : secr@genoscope.cnrs.fr - Web : www.genoscope.cnrs.fr)
COMMENT	This GSS is part of a random genomic sequencing program of thirteen yeast species: <i>Saccharomyces bayanus</i> var. <i>uvvarum</i> , <i>Saccharomyces exigua</i> , <i>Saccharomyces servazzii</i> , <i>Zygosaccharomyces rouxii</i> , <i>Saccharomyces kluveri</i> , <i>Kluyveromyces thermotolerans</i> , <i>Kluyveromyces lactis</i> var. <i>lactis</i> , <i>Kluyveromyces marxianus</i> var. <i>marxianus</i> , <i>Pichia angusta</i> , <i>Debaromyces hansenii</i> var. <i>hansenii</i> , <i>Pichia sorbophila</i> , <i>Candida tropicalis</i> and <i>Yarrowia lipolytica</i> . Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See 5 kb were prepared and both extremities were sequenced. See the other extremity of this sequence and for the sequence of the other extremity of this insert.
FEATURES	Location/Qualifiers
source	1. . 905
	/organism="Debaromyces hansenii"
	/mol-type="genomic DNA"
	/strain="CBS 767"

capture the induction and initiation/initial growth of the tuber in an in vitro system as described in Bachem et al. (Plant Journal, 1996). Small microtubers develop from axillary buds attached to stem explants when placed on a high sucrose medium (10%). Visible morphological changes occur synchronously at day five in the axillary buds. The first library, cSTA (11-20) consists of axillary buds harvested on days 1-3. This targets those genes involved in induction of the microtubers. The following libraries, cSTA (21-40) and cSTA (41-60), capture genes involved in tuber initiation and outgrowth. This library is noted as P3 in Tanksley lab notebooks.

ORIGIN

Query Match 1.6%; Score 103.6; DB 12; Length 781; Best Local Similarity 48.0%; Pred. No. 2.3e-08; Matches 360; Conservative 0; Mismatches 384; Indels 6; Gaps 2;

Db

4924 TTGGATGTTACACTAAAGAGCTTGTGAACTATGAGCTGGCTA 4803

QY 697 GTATATGTTGATCATGTTAATGGTAAAGTGTGTTCTTCTGAT 756

Db

4984 ACACATGTTGTTACATGAAATGAGGGATCAGTAATATTAGTT 5043

QY 757 TGCCTATGAGGATGTTGTAATATTTCGAAATAGAGCTTGTGACT 816

Db

5044 ATATGTGATGATCT 5062

QY 817 ATATGTTGAGATGTT 835

RESULT 4

Bi92632

LOCUS Bi92632 781 bp mRNA linear EST 10-MAR-2003

DEFINITION potato microtubers, in vitro-grown *Solanum tuberosum* cDNA clone cSTA6G17 5' end, mRNA sequence.

ACCESSION Bi92632

VERSION Bi92632.1 GI:16216672

KEYWORDS EST.

SOURCE *Solanum tuberosum* (potato)

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 781)

Van der Hoeven, R., Bezzerril, J., Bachem, C., Visser, R., Chiemingo, A., Karamycheva, S.A., Tsai, J., Van Aken, S., Utterback, T., Chiemingo, A., Bougri, O., Bueli, C.R., Ronning, C., Tanksley, S. and Baker, B. Generation of ESTs from in vitro grown microtubers (2001b)

JOURNAL Unpublished (2001)

COMMENT Contact: Robin Buij

The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Email: potato-array@tigr.org

This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
<http://genome.arizona.edu/orders/>

Seq primer: T3.

FEATURES

source

1. .781 Location/Qualifiers

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Bintje"

/db_xref="taxon:4113"

/clone="cSTA6G17"

/tissue_type="axillary buds of stem explants; growing

/sink_tubers"

/dev_stage="7, 8 and 10 days"

/lab_host="SOLR"

/clone_id="potato microtubers, in vitro grown"

/note="Vector: pBluescript SK(-); Site1: EcoRI; Site2: XbaI; Tissue supplied by Christian Bachem and Richard Visser (Department of Plant Breeding, Wageningen University, The Netherlands); sequencing by The Institute for Genomic Research. The cSTA libraries will attempt to

Db

4684 GGATTTGACCTTGTAGTTAGTCACCTGTTAGTCATGCTGTGACTATAGATT 4743

QY 458 ACATCTGATACCTCGAAGTCGTCATCTGTGATGACTATCCATAATTGTTACT 517

QY 4744 GACATAATAAGGTGTGAAATTAGGATGAACTAACAGTTAGACGTGGCTA 4803

Db 518 TACAGGAAAGAGT-TGATATGACTGGCGATACATGATCTTGATATGAGTT 576

QY 4804 TCTAATGCTCTATGTTACATCAATCAATTATGTTCTCTCTTAATCAGTAC 4863

Db 577 TTCCCTGCCGATATGCAATCASNCAVCCATATVYCTACCCACCTGGCTGAC 636

QY 4864 TTGGAGAAACATGTGTTATGAACTGTTCTATGGTTAAACAGTGGG 4923

Db 637 TATGAYTCYCGTAAATSTGTTGATTCCATCTGTTCTGTTCTGTTGAGGG 696

QY 4924 TTGGATGTTACACTAAAGAGCTTGTGAACTATGAGCTGGCTA 4803

Db 697 GTATATGTTGATCATGTTAATGGTAAAGTGTGTTCTTCTGAT 756

QY 4984 ACACATGTTGTTACATGAAATGAGGGATCAGTAATATTAGTT 5043

Db 757 TGCCTATGAGGATGTTGTAATATTTCGAAATAGAGCTTGTGACT 816

QY 5044 ATATGTGATGATCT 5062

Db 817 ATATGTTGAGATGTT 835

Db

4666 CAGACAAAGAAATTGATTAGACCTTTAGCTGTAGTCACGTGTTGTCATGCA 4665

QY 4667 TGAGAAATGATCTCAAGGTTGTCGAATACAGTGTGTTGCGCA 191

Db 192 AGGGTAACTCAGATTTGGCTAGATTTAGTCACATTCTCTGGCTAAAT 251

QY 4726 TGTGATATAGATTTGCAATATAGTTGTTGATTTAGGATGACATTCAACATT 4785

Db 252 TGCATTTGTTGCTCTTCTATCTATGTTGTCATGCGCTCTTCATCAGT 311

QY 4786 AGACGTGAGTGGCCATTAATGCCCTATTAATCTCAATTAAATTATGCTT 4845

Db 312 GGACATTAAGATGCTTCTGCTGTTGTTGAGGAAAGTCTATATGGCGAAC 371

QY 4846 TCCTCTTAATCTGACTCTTGAG---AAACCATTTGGTATGAACTTCTG 4902

Db 372 ACCTGGTTGTTGCTCAGGGGAGCTAGTGTGCTTGTGTTGCTGAGCT 431

QY 4903 CTATGGTTAACACAGTCGGTTGGATGGTATCACATCAAAAGGATGGAGA 4962

Db 432 CTATGGCTAACACCTCTCTCAGCTGCTGTTGGAGGTTCGACAGTAATCAGA 491

QY 4963 CATTGGTTTACTCAAGTTACACATGATGGTTTACATTGAAATGAGAGG 502

Db 492 GTTGGCATGACTCGTAGTGGAGCTACTCTGTGTTATGGCATCTGCACAG 551

QY 5023 ATCAGTAAATAATTGGTTATAGTGTGATGATCTTATGTTGGAGTTACAAA 5082

Db 552 TCAATCTCATTTGTTGTTGATGATGTTACCGGATATGCAAGA 611

QY 5083 AGTTATGATAATTGTTGTTGCTATGAGGATCATTTGAGTTAAGTGTGTC 5142

Db 612 TGTATACCGATTTGAGCACATCTTAAACCTTCAGACTTAAGACCTTGCG 671

QY 5143 ATATCAAATCTGTTAGTGAATTCGTAACACGAACTCTGTTATGGTATTC 5202

Db 672 ATTGAGTATTCTGGATGTTGCTCAGTGTGTTATGCTCA 731

QY 5203 AGAAATTCTCAAGAAATTCTAGGA 5232

Db 732 ACCGAACTGCTTACGATCTCTGAGGA 761

LOCUS	AT813079	569 bp	mRNA	linear	EST 08-JUL-1999
DEFINITION	21BII Pine Lambda zap Xylem library Pinus taeda cDNA, mRNA sequence.				
ACCESSION	AI813079				
VERSION	AI813079.1				
KEYWORDS	EST.				
SOURCE	Pinus taeda (loblolly pine)				
ORGANISM	Pinus taeda; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus, 1 (bases 1 to 569)				
REFERENCE	Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.				
AUTHORS	Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.				
TITLE	The Pine Gene Discovery Project				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Ross Whetten Forest Biotechnology Group North Carolina State University Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh, NC, 27695-8008 Tel: 919-515-7800 Fax: 919-515-7801 Email: rossw@unity.ncsu.edu				
FEATURES	Seq primer: T3.				
source	Location/Qualifiers				
1. 569					
/organism="Pinus taeda"					
/mol_type="mRNA"					
/db_xref="taxon:3352"					
/tissue_type="differentiating xylem"					
/clone lib="Pine Lambda Zap Xylem library"					
/note="Vector: Lambda Zap; Site_1: BcoRI; Site_2: XbaI"					
Differentiating xylem was collected from the main stem of a 35-year old loblolly pine tree harvested during the growing season. RNA isolation and library preparation followed the methods of Allona et al., PNAS 95:9693-8, 1998"					
ORIGIN					
Query Match	1.5%; Score 97.8; DB 9; Length 569;				
Best Local Similarity	48.5%; Pred. No. 2.6e-07;				
Matches	270; Conservative 0; Mismatches 287; Indels 0; Gaps 0;				
OY	4730 ACTATAAGATATGACAATATAGGTGTGAATGACAATTAGAC 4789				
Db	13 ACAATTAGATATGATGATGTTCTAAATGAGCATACATCAATTGAG 72				
QY	4790 GTCGAGATCGGGTATCTAACTCTTAACTTCAATCCATTCTCT 4849				
Db	73 GTCAAATCTGCTTCTAAATGGATTAAGAAGAGTATTGGTACCTGAA 132				
QY	4850 CCTAAATCAGTACCTTGAGAAACACATTGTTGTTAGAACGTTCTGTCTATGG 4909				
Db	133 GGGTGTGCAACAGGGACGGACATCTGGTGGAGTTAGAAAGCATGTGT 192				
QY	4910 TAAACACAGTGGGTTGGAATGGTATCACATCAAAAGAGTATGGAGACATGGT 4969				
Db	193 TGAACACGACCAAGGGCATGGATGAAAGTACCTCTCTCTGAAATGGA 252				
QY	4970 TTACTCAAGTTTACACATGAGGTATTACATGATAATGAGGAGTCGTA 5029				
Db	253 TTGTGAAAGCAAAATGATCCTAATCTGATGTTGAAAGATGAAAGGAAATGT 312				
QY	5030 ATATATTAGTTATATGTTGATGATATCTTATGCTGGAGTCACAAAGTATT 5089				
Db	313 GCTTAAATCTTGTGATGATGATGATCTTAAACGGATGTTGCTACAGTTATT 372				
QY	5090 GATAATTGTTGGAATGAGATCATTTGAGTAAGTGTGGAAATATCA 5149				
Db	373 GAGAATTAAGTCAAGTGTCAAGGATTTGAATGAAAGATGGTAAATGCA 432				
QY	5150 ATTAATCTGGATGAAATTGCTAAACCGAACTGGTAAATTAACTCAAGA 5209				
Db	433 TACTGTTAGAATTGGAGTATGGAGAGAACCTGGTAAACTTGTACCTCAAACAA 492				
FEATURES	source				
	Location/Qualifiers				
1.	1..1035				
	/organism="Debaromyces hansenii"				
	/mol_type="genomic DNA"				
	/strain="CBS 767"				
	/variety="hansenii"				
	/db_xref="txon:4959"				
	/clone="BC00005F10"				
	/clone_id="BC00AA"				
	/note="end : T7"				
	<1. ->0.35				
	/note="part of putative transposable element"				
	/evidence="non_experimental"				
	ORIGIN				

COMMENT	Other_GSSs: MBEDX18TRC Contact: Chris Town TIGR 971 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-888-3523 Fax: 301-888-0208 Email: cdtown@tigr.org Seq primer: CAGGAAACGCTATGACC Class: BAC ends.
FEATURES	source 1. .766 /organism="Medicago truncatula" /mol_type="genomic DNA" /cultivar="genotype Al7" /db_xref="taxon:3880" /clone="35D12" /note="Vector: pBelobAC1; Site_1: HindIII; Site_2: HindIII; Cook, D.R. and Kim, D.J., unpublished"
ORIGIN	Query Match Best Local Similarity 1.4%; Score 92.8; DB 29; Length 766; Matches 286; Conservative 0; Mismatches 322; Indels 0; Gaps 0; Qy 4725 TGTGACTATAGAAGTATTGACAATAATGGTTGGAATTGAAATCAATT 4784 73 TGGATACATTGAACTCTGTGAGGTGGCTACAGTACAAATAGGGAAATTGGATT 132 Qy 4785 TAGAGTCGAGTCGGTAACTTAATGCTCTATACATCAATCAATTATGCT 4844 133 TAGATTCATACAGCATCTTAATGAAATTCTGTAGAGAAATTATGTGCTAAC 192 Qy 4845 TCCCTCTAAATCAGTACCTTGAGAAACCATGTGTTGTTGAAACCTCTCT 4904 193 CTGCGTGTGTTTGTAAAGGGAGGAGAACAGGTGATAGCTCATAGCTTGT 252 Db 4905 AAGGTTAAACAGTGGTTGGATGTTACACTAAAGATGGAAAGACA 4964 253 ATGGCTGAAACAGGCCCTAGAGCTTACAAATTATCCATTCTTAACTCTTAACT 312 Qy 4965 TGGTTTACTCAACTTACACATGATGGTTATTGACATGAAATGAGGCGAT 5024 313 AAGGTTCAAAAGGGTGAATGCAACTCTTAATGAAAGCTGTGATGGT 372 Db 5025 CAGTAAATTTAGGTTATGATGATTAATCTATGTTGGAGTTCAAAAG 5084 373 GTTCTCTAAATGTCCTTGTGATGACCTGCTAGAACGAAATCATCAACAG 432 Db 5085 TATGATAATTGAGGCAATTGAGATCATTTGAGTTAAAGTGTGGTCAA 5144 Qy 433 AAGTACATCACTTATGGAGGATGAAACCGATTTGAGATCTAGCTAGGGAA 492 Db 5145 TATCAAATCTCTGATGATTGAAACGAATCTGGTATATTCTCAAG 5204 493 TGACTATTCCTGCTGTGAGTGCACTGTGAGATTTCATG 552 Qy 5205 AAAATTCTCAAGAAATTCTAGGATTCAACTATGACTCATATGAAACA 5264 553 AGAAGAATGTCGATGAGTTGAGAACTTAATGGAAAGGTGCAATCTGCTCAA 612 Db 5265 TACCTCTGAACTTCGAGAAATATGAAAGGTGCAATAATTGAAAGTAACT 5324 613 CTCCTTGGTGGATTGAGCTCAAGGAGATGACCTGAACTGATGCT 672 Qy 5325 CAGGAAAT 5332 Db 673 CTCATTT 680 RESULT 9 CG93798 CG93798 LOCUS CG93798
DEFINITION	MBECA46TF mth2 Medicago truncatula genomic clone 24G19, genomic survey sequence.
ACCESSION	CG337798
VERSION	CG337798.1 GI:39809821
KEYWORDS	GSS:
SOURCE	Medicago truncatula (barrel medic)
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago
REFERENCE	1 (bases 1 to 866)
AUTHORS	Town,C.D., Shetty,J., Koo,H. and Feldblyum,T.F.
TITLE	Sequencing of BAC ends from Medicago truncatula
JOURNAL	Unpublished (2003)
COMMENT	Other_GSSs: MBECA46TR Contact: Chris Town TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-888-3523 Fax: 301-888-0208 Email: cdtown@tigr.org Seq primer: TGTAAACGACGGCCAGT Class: BAC ends.
ORIGIN	Query Match Best Local Similarity 1.4%; Score 91.4; DB 29; Length 866; Matches 272; Conservative 0; Mismatches 301; Indels 0; Gaps 0; Qy 4725 TGTGACTATAGAAGTATTGACAATAATGGTTGGAATTGAAATCAATT 4784 73 TGGATACATTGAACTCTGTGAGGTGGCTACAGTACAAATAGGGAAATTGGATT 132 Qy 4785 TAGAGTCGAGTCGGTAACTTAATGCTCTATACATCAATCAATTATGCT 4904 133 TAGATTCATACAGCATCTTAATGAAATTCTGTAGAGAAATTATGTGCTAAC 192 Qy 4845 TCCCTCTAAATCAGTACCTTGAGAAACCATGTGTTGTTGAAACCTCTCT 4904 193 CTGCGTGTGTTTGTAAAGGGAGGAGAACAGGTGATAGCTCATAGCTTGT 252 Db 4905 ATGGCTGAAACAGGCCCTAGAGCTTACAAAGAGTATTGGAGACA 4964 253 ATGGCTGAAACAGGCCCTAGAGCTTACAAAGAGTATTGGAGACA 312 Qy 4965 TGGTTTACTCAACTTACACATGATGGTTATTGACATGAAATGAGGCGAT 5024 313 AAGGTTCAAAAGGGTGAATGCAACTCTTAATGAAAGCTGTGATGGT 372 Db 5025 CAGTAAATTTAGGTTATGATGATTAATCTATGTTGGAGTTCAAAAG 5084 5025 CAGTAAATTTAGGTTATGATGATTAATCTATGTTGGAGTTCAAAAG 5084 Db 5145 TATCAAATCTCTGATGATTGAAACGAATCTGGTATATTCTCAAG 5204 313 AAGGTTCAAAAGGGTGAATGCAACTCTTAATGAAAGCTGTGATGGT 372 Db 5265 TACCTCTGAACTTCGAGAAATATGAAAGGTGCAATAATTGAAAGTAACT 5324 373 GTCCTTAATGTCCTTGTGATGACCTGCTAGAACGAACTGTTCAAAAG 432 Qy 5325 TATGATAATTGAGGCAATTGAGATCATTTGAGTTAAAGTGTGGTCAA 5144 433 AAGTACATCACTTATGGAGGATGAAACCGATTTGAGATCTAGCTAGGGAA 492 Db 5493 TGACTATTCCTGCTGTGAGTGCACTGTGAGATTTCATG 552 Qy 5145 TATCAAATCTCTGATGATTGAAACGAATCTGGTATATTCTCAAG 5204 313 AAGGTTCAAAAGGGTGAATGCAACTCTTAATGAAAGCTGTGATGGT 372 Db 5525 CAGTAAATTTAGGTTATGATGATTAATCTATGTTGGAGTTCAAAAG 5084 5025 CAGTAAATTTAGGTTATGATGATTAATCTATGTTGGAGTTCAAAAG 5084 Db 5145 TATCAAATCTCTGATGATTGAAACGAATCTGGTATATTCTCAAG 5204 493 TGACTATTCCTGCTGTGAGTGCACTGTGAGATTTCATG 552 Qy 5205 AAAATTCTCAAGAAATTCTAGGATTCAACTATGACTCATATGAAACA 5264

Db	553	AGAAGATCTCATGAGTTGAGAAGTTAAATGAAAGCTGCAAATCTGCTCAA	612	Db	373	GTTCCTTAATGGTCTTGTGAGTGTGAGCTGCTGAGTACAGCAATACAAAG	432
Qy	5265	TACCCGGATTCGGAATGACAATATGAAAGG 5297		Qy	5085	TATGATAATTGTTGGATCAATGAGATATTTGAGTTAAGTGTGGTGAAG	5144
Db	613	CTCCTTGGTGGATATGAAACTCTCAAGG 645		Db	433	AAGTACATCACTATGGAGGAGTGAACAAACCGAGTTGAGTGTCTAGTAGGGAA	492
				Qy	5145	TACCAATTATCTGGTATGTTCTGAAACCGAACTGGTTATTTTATTCAG	5204
				Db	493	TGAACATTCTCTGGCTGGAGTCACATCTGAGGAGTTGGATCAAGG 552	552
RESULT 10	CG973377	CG973377	867 bp	DNA	linear	GSS 15-DEC-2003	
DEFINITION	MBED30TFC	mtb2	Medicago truncatula	genomic clone	35E12,	genomic	
REFERENCE	CG973377						
AUTHORS	Town,C.D., Shetty,J., Koo,H. and Feldblum,T.F.						
TITLE	Sequencing of BAC ends from Medicago truncatula						
JOURNAL	Unpublished (2003)						
COMMENT	Other_GSS: MBEV30TRC						
ORGANISM	Medicago truncatula (barrel medic)						
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Spermatophytina; Magnoliophyta; eudicots; rosids; euroids; I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago							
(bases 1 to 867)							
Town,C.D., Shetty,J., Koo,H. and Feldblum,T.F.							
Sequencing of BAC ends from Medicago truncatula							
Unpublished (2003)							
Other_GSS: MBEV30TRC							
Contact: Chris Town							
TIGR							
9712 Medical Center Drive, Rockville, MD 20850, USA.							
Tel: 301-838-3523							
Fax: 301-838-0208							
Email: cdtown@tigr.org							
Seq primer: CAGGAACAGCTATGACC							
Class: BAC ends							
FEATURES							
SOURCE							
ORIGIN							
Query Match							
Best local Similarity	1.4%	Score 91.4;	DB 29;	Length 867;			
Matches	272;	Conservative	47.5%;	Pred. No. 3.4e-06;			
Qy	4725	TTCGACTATAGAGATATGACAATAATGGTGTGGAATAGGATGACATT	4784				
Db	73	TGGACTACTATGGAATCTGAGGTTGGCGAGTACAAATTAAGTGGCATT	132				
Qy	4785	TAGAGCTGAGTCGGGTCTAAATGCCCTATTAATGCAATTATGCAATTATGCT	4844				
Db	133	TAGATTCAATACAGCATCTTAAATGGAATCTGGTGTGTCAC	192				
Qy	4845	TTCCTCTAAATCAGTACCTTGAAGAAACCTGTGTGTATGAACTGTCCT	4904				
Db	193	CTGGCTGGTTTGGTAAAGGGAGGAGACAGGGTGTAGTCATAAGCTGCT	252				
Qy	4905	ATGGCTAAACAGCTGGGTGGATGATCACTATCAAGAGATGGAGACA	4964				
Db	253	ATGGCTGAACAGGCCCTAGAGGTGGACACAAATTAAATTCACCTCTTAAC	312				
Qy	4965	TTCGTTTACTCAAGTTTACACATGATGTTTACATGAGTAAATGAAAGGG	5024				
Db	313	AAGGTTCAACAGGGTGAATGTCACCTTATGAGGCTGTGGTGGC	372				
Qy	5025	CAGTATATATTAGTTATGTCAGATATCTTATGGTGTGAGTCACAAAAG	5084				
Db	193	CTGGTGGTTTGGTAAAGGGAGGAGACAGGTGTAGTCATAAGCTGCT	252				
Qy	4725	TTCGACTATAGATTGACAATAATGGTGTGAAATGACATCAACATT	4784				
Db	73	TGGACTACTATGGAATCTGAGGTTGGCGAGTACAAATTAAGTGGCATT	132				
Qy	4785	TAGAGCTGAGTCGGGTCTAAATGCCCTATTAATGCAATTATGCAATTATGCT	4844				
Db	133	TAGATTCAATACAGCATCTTAAATGGAATCTGGTGTGAAATTTGTCAC	192				
Qy	4845	TTCCTCTAAATCAGTACCTTGAAGAAACCTGTGTGTATGAACTGTCCT	4904				
Db	193	CTGGCTGGTTTGGTAAAGGGAGGAGACAGGTGTAGTCATAAGCTGCT	252				

Qy	4905	ATGGGTTAAACAGTCGGGTTGGAAATGGTCACTATCAAAAGAGTATGGAAAGCA	4964	Qy	4483	TAATGATCCGGATGGATAAAGTCATGATGTCAGCTAGAGAAATTAGATCAAAGA	4542
Db	253	ATGGGTTGAAACAGGGCCCTAGACTGTTACACATGATGTTATTGAAATTCACCTCTATC	312	Db	118	TAACCTGAGCTTGGAAGGATGCTATGGAGGATCAACTCCTTGAGAAACTCG	177
Qy	4965	TGGGTTTACTCAAGTTTACACATGATGTTATTGAAATGAGAGGAT	5024	Qy	4543	TGTTTACGAAAGTTCGAATCCACCTGAGCCATATCTATCGGTGGTACA	4602
Db	313	AGGAGATCAAAAGGAGTGAATAATGATGCAACTCTTATGAGAGCTGTTG	372	Db	178	TACTGAAACTTGTTGAGATTACACCAAGTGAAGCTGAGGATTA	237
Qy	5025	CAGGATATATTAGTTATGATGATGAACTCTATGTTGAGGTTACAGAAG	5084	Qy	4603	TACTGAGAAATTGATCTTCGAAGGGTGTGAGGAATCAAGTGTGTTGCG	4662
Db	373	GTTCTTAATTGTCCTTGATGTTGAGCACTGCTAGTCAAGAAG	432	Db	238	CAAGTCAAATTTCAGCT---GATGGCACTATTGAGAGATAAAAGCAAG	294
Qy	5085	TATGATATTGTTGTTGAACTGAGGATATTGAGTTAAGTGTGTTGAA	5144	Qy	4663	CAACGACAAAGGAATTGGATTGACATAATGAGCTTGTAGGAA	354
Db	433	AGTACATCACTATGAGGGAGTGAACACAGTTGAGGCTAGTGGCAA	492	Db	295	CAASGTTACATCAATGAGGCTTAAATTGAGCTTGTAGGAA	414
Qy	5145	TATCAATTACTCTGGTATGAACTGAAACCGAACTCTGTTTATTTATCTCAAG	5204	Qy	4723	TCTGTGCACTATAAGATTGACATAATGAGCTTGTAGGAA	4722
Db	493	TGAACTATTCTCTGGCTGGAGTCATCACTGAGGTTTGATGATCAG	552	Db	35	ACTCACACAGTCAGCTGTGTTATCTTCATATGCCATTACATCA	414
Qy	5205	AAAATTTCAGGAAATACCTTAACTGATGACTCATGGAAACAA	5264	Qy	4783	TTAGAGTCAGTCGGGAGTCAATCCTTATTACTCATCAATTCAATTATG	4842
Db	553	AGAAGATGCTCATGAGTTGAGAACTTAAATGAAAGCTGCAATCTGCTCAA	612	Db	415	ACTGTGTCATAATGCTCTTCATG---GTOAACTCAGGAGTGTGACATGT	471
Db	613	CTCCCTTGTTGGATGGATTGAGCTCAAGCTCAAGG	645	Qy	4843	CTTCTCTTAATGATGACTTGTAGGAAACCATGTTGTTTATTAAACCTCT	4902
RESULT 12				Db	472	GATCCCTCTGTCATCAACCAATCAACAAATCAAGTGTGAGCTCAAAGCT	531
CG28381	CG28381	MBEIG3TR mth2 Medicago truncatula	940 bp	Qy	4903	CTATGGTTAACAGTCGGGTTGGAAAGTGTACACTATCAAAAGGTATGGAA	4962
DEFINITION		Medicago truncatula genomic clone 61H3, genomic	DNA	Db	532	CTATGGACTTAATGGCTAGTAGAAGTGTAGGAA	591
ACCESSION	CG928381		linear	Qy	4963	CATGGTTTACTCAAGTTTACAAATGGTTATTACATGATCTACTCTTGGC	591
VERSION	CG928381.1		GSS	Db	592	TCAAGACTATTTACAGTCATGATCTTCCTGTCAGAAACTGAA---TC	648
KEYWORDS	GSS			Qy	5023	ATCAGTATAATTAGGTTATGTTGATGATTCCTATGGTGGAGTACAAA	5082
SOURCE	Medicago truncatula (barrel medic)			Db	649	ATCATTCACTGTTCTGTTGTTGATGATCATCTCTGCTGGAGTCTATAAG	708
ORGANISM	Medicago truncatula			Qy	5083	AGTTATGTTATTGTTGATCACTATGGAGATATTGAGTTAAGTGTGAGA	5142
BACTERIA	Bacillus			Db	709	TGATTGTCATCACTGATCTGTTGAGTACATTCATTGTTGAGTACAAA	768
SPERMATOPHYTE	Streptophyta			Qy	5143	AAATCAAAATATCTGGTTGTTGATTGTTAAACCGATCTGTTTAAATTATC	5202
RODENTS	Magnoliophyta			Db	769	ATGAAATACCTCTGGTTGAGGTTGCTCATTCATTGGGATTCTGTTGAG	828
EUROPSIDS	Eudicots			Qy	5203	AGAAMATTTCAGGAAATCTAGGTTAACCTGAGATGACTC	5250
Medicago	Fabaceae			Db	829	AAGGAATATTGTCAGACTTACTCTGTTGAGGATTAGGTC	876
REFERENCE	1 (bases 1 to 940)						
AUTHORS	Town,C.D., Shetty,J., Koo,H. and Felddiium,T.F.						
JOURNAL	Sequencing of BAC ends from <i>Medicago truncatula</i>						
COMMENT	Unpublished (2003)						
CONTACT	Other_GSS: MBEIG43TRB						
TIGR	Chris Town						
	9712 Medical Center Drive, Rockville, MD 20850, USA.						
	Tel: 301-888-3523						
	Fax: 301-888-0208						
	Email: cdtown@tigr.org						
	Class: BAC ends						
FEATURES	Seq primer: TGTAAACGGGGCCAGT						
SOURCE	Location/Oulifiers						
	1. .940						
	/organism="Medicago truncatula"						
	/mol_type="genomic DNA"						
	/cult_iva="genotype A17"						
	/db_xref="taxon:3880"						
	/clone_id="61H3"						
	/clone.lib="mth2"						
	/note="Vector: pBelobAC1; Site_1: HindIII; Site_2:						
ORIGIN	HindIII; Cook, D.R. and Kim, D.J, unpublished"						
Query Match	1.4%; Score 89.6; DB 29; length 940;						
Best Local Similarity	47.5%; Pred. No. 7e-06;						
Matches	365; Conservative 0; Mismatches 394; Indels 9; Caps 3;						
RESULT 13							
CG821291	CG821291		803 bp				
DEFINITION	SORBA07V Large Insert Soybean Genbank		DNA				
ACCESSION	B4J01-MTP07011		linear				
VERSION	CG821291.1		GSS				
KEYWORDS	Glycine max (soybean)						
SOURCE	Glycine max						
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Fabaceae; Papilionoideae; Phaseoleae; Glycine.						
REFERENCE	1 (bases 1 to 803)						
AUTHORS	Shultz,J., Meksem,K., Shetty,J., Town,C.D., Koo,H., Potter,J., Waterfield,K., Zhang,H., Wu,C. and Lightfoot,D.A.						
TITLE	End sequencing of BACs comprising a provisional minimal tiling path from a fingerprint physical map of soybean (Glycine max) cultivar Forrest						

JOURNAL		COMMENT
Unpublished (2003)		Other_GSS: SOYBA90TH
Contact: Chris Town, J. L. Shultz and D. A. Lightfoot		
The Center of Excellence in Soybean Research, Teaching and Outreach, Southern Illinois University at Carbondale and Plant Genomics, The Institute for Genomic Research, Rockville, MD 20850, USA and 9712 Medical Center Drive, Rockville, MD 20850, USA		Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415, USA
Re: 618 453 1779		Email: ga408@siu.edu, Jlshultz@siu.edu, cdtown@tigr.org (URL: http://bioinformatics.siu.edu)
4.33 ACCAACCGAGATTTGTCGATGAGGAGTCAGCTTGCAACCTCGCAGGCTCT	433	Clones approximating a minimum tiling path were re-arrayed from the library master plates prior to sequencing.
For purposes of clone identification each clone name is a concatenation of the original clone location and its new location in the re-arrayed sequencing plates.		Seg primer: GTAAATACGACTCACTATAGGC
Class: BAC ends.		
Location/Qualifiers		
1. .803		
/organism="Glycine max"		
/mol_type="genomic DNA"		
/cultivar="Forrest"		
/db_xref="taxon:3847"		
/clone="B47J01:MP101"		
/clone_lib="LargeInsertSoybeanGenLib"		
/note="Organ: Leaves; Vector: pCIB04541 (pBLOBAC11 EcoRI clones); Site_1: BetX1; Soybean (Glycine max (L.) Merr.) cv. Forrest seeds were grown in greenhouse for fourteen days. Nuclei were isolated and embedded in agarose, restriction digested with Hind III BamHI or EcoRI, large size DNA fragments were ligated in vector pV41 (pCD0541) and electro transformed in DH10a cells. About 90,000 clones from BAC libraries were fingerprinted with HindIII and Hae III. Version 2 (automatic build) Contigs were examined to find the best non redundant tile path through the contigs representing about 13,000 clones. The clones were end sequenced."		
ORIGIN		
Query Match		
Best Local Similarity 1.4%; Score 89.4; DB 29; Length 803; Matches 343; Conservative 0; Mismatches 401; Indels 3; Gaps 1;		
QY 4483 TATGATTCGGGATGGATAAAGTCATGATGCTGAACTAGAGAAATTAGTCAAAGA 4542		
Db 790 TGTATCCGGATGGCACAGCCATGATGATGAAATACGGCTTGGACATAAG 731		
QY 4543 TGTATGGAGAGTCGAATTCACCGGNGTGAGCTATATCTATGGGTGGTACA 4602		
Db 730 TACTGGAGCTAGTCCTCTCCAGGCAGAACTGTGTGTCGGCTCA 671		
QY 4603 TACTGAGAAATTGATCTCAAAGGTGTTCGAATACGTTGTCATGG 4662		
Db 670 TGTCTAAAGTGCGGCCATAGACAGATTGATGACCTTTAGTGTGTCATGAGTAC-- 614		
QY 4663 CAAAGACAAAGGAATTGATGACCTTTAGTGTGTCACCTGTATAGA 4722		
Db 613 CAAAGGTATAACCGATTAGACCTAGATGAGTACCTTCCTGCTCTAA 554		
QY 4723 TCTGTGACTATAAGATTAGGACATAATAGTTGTGAATTAGGAATGACATCAACA 4782		
Db 553 AATTACTCTATTCGAGCTTCCTGCATAGCTGCCATCGCTCATGGCACTGATCA 494		
QY 4783 TTAGAGTCGGTGGCTATCTAAATGCTCTATACTCTTAAATCCATTATGT 4842		
QY 493 ATGGATTAATAGCATTCCTACTGGAAATTAGAAGGGAGTTATGAGCA 434		
Db 4843 CTTCCTCTAAATCAGCTTGAAGAAACATGTCGTTATGAAACGTCGT 4902		
Db 4.33 ACCAACCGAGATTTGTCGATGAGGAGTCAGCTTGCAACCTCGCAGGCTCT	433	Clones approximating a minimum tiling path were re-arrayed from the library master plates prior to sequencing.
For purposes of clone identification each clone name is a concatenation of the original clone location and its new location in the re-arrayed sequencing plates.		Seg primer: GTAAATACGACTCACTATAGGC
Class: BAC ends.		
Location/Qualifiers		
1. .803		
/organism="Glycine max"		
/mol_type="genomic DNA"		
/cultivar="Forrest"		
/db_xref="taxon:3847"		
/clone="B47J01:MP101"		
/clone_lib="LargeInsertSoybeanGenLib"		
/note="Organ: Leaves; Vector: pCIB04541 (pBLOBAC11 EcoRI clones); Site_1: BetX1; Soybean (Glycine max (L.) Merr.) cv. Forrest seeds were grown in greenhouse for fourteen days. Nuclei were isolated and embedded in agarose, restriction digested with Hind III BamHI or EcoRI, large size DNA fragments were ligated in vector pV41 (pCD0541) and electro transformed in DH10a cells. About 90,000 clones from BAC libraries were fingerprinted with HindIII and Hae III. Version 2 (automatic build) Contigs were examined to find the best non redundant tile path through the contigs representing about 13,000 clones. The clones were end sequenced."		
ORIGIN		
Query Match		
Best Local Similarity 1.4%; Score 89.4; DB 29; Length 803; Matches 343; Conservative 0; Mismatches 401; Indels 3; Gaps 1;		
QY 4903 CTATGGTTAACAGTCGGGTTGGAATGGTACACTACAGAGATATGGAGA 4962		
Db 373 TTATGACTCAACATCCACAGCTGGTTGGAATTCAGCTACTTGTCAAGGC 314		
QY 4963 CATTGTTTCTACGTTACAGATGTTTACACATGTTTATTACATGAAATGAGGG 5022		
Db 313 TTGAGGATGAGGAGTGAGCACCATTCAGTTTATGTCAGTCACTCATGAG 254		
QY 5023 ATCAGAATATAATTAGGTATATGTTGATATCTATGTTGGAGGTACAAA 5082		
Db 253 TAGATGTGTTACTGTGGCTATGATGATAGATAGTCACAGGAAATGAGGA 194		
QY 5083 AGTATTGATTTGTTGTTGCAATTGAGATCATTTGAGTTAAGTGTGTTGTA 5142		
Db 193 AAAATGCTCACTGAAAGACATTATAGTCATTTAGCATTAAGCTAAGGTGTGGA 134		
QY 5143 ATATAAAATTATCTGGTATGAAATTCTGAAACCGATCTGGTTATTTATCTA 5202		
Db 133 ACTTAATATTTCTGAAATTACTTAA 5229		
QY 5203 AGAAAATTCTCAGAAATTACTTAA 5229		
Db 73 AGAAGATGATCGTTAGACATACTTAA 47		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
QY 94943796 CG943795 926 bp DNA linear GSS 15-DEC-2003		
LOCUS		
DEFINITION		
MBEMC57TR mth2 Medicago truncatula genomic clone 84U17, genomic survey sequence.		
ACCESSION		
CG943796		
VERSION		
CG943795.1		
KEYWORDS		
GSS:		
SOURCE		
ORGANISM		
Medicago truncatula (barrel medic)		
Buarycca; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.		
1 (bases 1 to 926)		
Town, C.D., Shetty, J., Koo, H. and Feldblum, T.F.		
Sequencing of BAC ends from Medicago truncatula		
Unpublished (2003)		
Other_GSS: MBEMC57TR		
CONTACT: Chris Town		
TIGR		
9712 Medical Center Drive, Rockville, MD 20850, USA.		
Tel: 301-838-3523		
Fax: 301-838-0208		
Email: cdtown@tigr.org		
Seg primer: GTAAATACGACTGCCAGT		
Class: BAC ends.		
Location/Qualifiers		
1. .926		
/organism="Medicago truncatula"		
/mol_type="genomic DNA"		
/cultivar="genotype A17"		
/db_xref="taxon:3880"		
/clone="B47J17"		
/clone_lib="mth2"		
/note="Vector: pBLOBAC11; Site_1: HindIII; Site_2: HindIII; Cook, D.R. and Kim, D.J., unpublished"		
ORIGIN		
Query Match		
Best Local Similarity 1.4%; Score 89.2; DB 29; Length 926; Matches 324; Conservative 0; Mismatches 343; Indels 9; Gaps 2;		
QY 4625 AAAGGTTGTCGGAAATCGTGTGTTGTCATGCCAACAGAACAAATG 4684		
Db 18 AAAGTGTGCAAAACAAACGCTGTTAGCTCAAGGTACACAAACAGAGSTATT 77		
QY 4685 GATTATGACCTTTAGTGTGTTGTCACCTGTTAGATGTCGTGACTATAGATTG 4744		

78 GACTATGA-----TGGACCTTGGACCACTTGGAGGTAGAACCCATCAGAATCTT 131
 /mol type="genomic DNA"
 strain="Miyakojima MG-20"
 variety="japonicus"
 /db xref="taxon:34305"
 /clone="LJ59017_sfi"
 /clone lib="genomic TAC library"
 /note="VECTOR: pYI-TAC7-synonym: Lotus japonicus"

QY 4745 ACAATATAGTTGTAATAGGAAAGACATTAGACGGTCGAGTGGTAT 4804
 /organism="Lotus japonicus"

Db 132 CTGGCTATGTCACATAAAGCATTAACTTTCAATGGATGTTGAAAGGATT 191
 /organism="Lotus japonicus"

QY 4805 CTAATGCCCTATTAATCTCATCAATCCAAATTATGTCCTTCCTCTCTAATCAGTACCT 4864
 /organism="Lotus japonicus"

Db 192 TTAATGGTTCTATGAGAAGGTATGAGTCACCTGGATGTAACCAAC 251
 /organism="Lotus japonicus"

QY 4865 TTGAAGAAACCATTTGTTGTTATTGAAAGGTCTGCTATGGTTAACAGTCGGGT 4924
 /organism="Lotus japonicus"

Db 252 GAAACCAAACTCATTTCAATTAAACAAAGGCTTATGGCTAAACAGCTCCT 311
 /organism="Lotus japonicus"

QY 4925 TTGGAAATGTTCACCTATCAAAGAGTAACTGGAAAGACATGGTTTACTCAGTTA 4984
 /organism="Lotus japonicus"

Db 312 AGAGCTTGGATGACAGACTTAGCACATTAAATTGAAATGGTTAACAGGAA 371
 /organism="Lotus japonicus"

QY 4985 CACAGTGTGTTTACATTGATGAGAGGGTCACTGATATATTGGTTA 5044
 /organism="Lotus japonicus"

Db 372 ATTGATACACACTTTAGAAACACATATACTGACTTGCTTATGTCAG --TA 428
 /organism="Lotus japonicus"

QY 5045 TATGTTGATGATTTTATGGTGGAGTTCACAAAGTATGATAATTGIGGAT 5104
 /organism="Lotus japonicus"

Db 429 TATGTTGATGACATTTTGGTTACTAAATAAATGTCAGGAGATTCCAC 488
 /organism="Lotus japonicus"

QY 5105 CAATGGAGATCATTGTAAGTTAAAGTTGGTGAATATCAATTACTGGATT 5164
 /organism="Lotus japonicus"

Db 489 CTTATCCTAACGTGAATTGAGATGACATGAGCTGGGAGACTGGTTTCCTGGCTA 548
 /organism="Lotus japonicus"

QY 5165 GAATTCTGTAACACCAATCGTTTATTTCTCAAGRAAATTTCAGAATT 5224
 /organism="Lotus japonicus"

Db 549 CAATCAACACATTCAATGGATTATACTGAGATATCAAGACATT 608
 /organism="Lotus japonicus"

QY 5225 CTTAAGGATTCAACTAGTACTCTATGGAAACATACCTGGATTCCGGATGAC 5284
 /organism="Lotus japonicus"

Db 609 CTTAAAGAATAATAATGAGGACAAATCATGACGACCCATTGATCCATCTCA 668
 /organism="Lotus japonicus"

QY 5285 AAATATGAAAGGTG 5300
 /organism="Lotus japonicus"

Db 669 AGTTTGTAGATAGGATC 684
 /organism="Lotus japonicus"

ORIGIN

Query Match 1.4%; Score 88.6; DB 29; Length 490;
 Best Local Similarity 53.1%; Pred No. 1.2e-05;
 Matches 212; Conservative 0; Mismatches 184; Indels 3; Gaps 1;
 Matche 212; Conservati 0; Mismatches 184; Indels 3; Gaps 1;

QY 4888 ATGAAACGTTCTGCTATGGTTAACAGTCGGTCTGGAAGGGTACACATATCAA 4944
 /organism="Lotus japonicus"

Db 444 ATTAAAGGCTATGATGCTTAACACAGCACCAAGTCTTGTACAGCGAGATGA 385
 /organism="Lotus japonicus"

QY 4948 AGAGTATGGAGACATGGTTACTCAGTTACAAAGTGGTTTACATGAGGGTTTACAT 5007
 /organism="Lotus japonicus"

Db 384 TACACATTAGAAGGCTAGGCTTGTTGAAAGGCTA--AGTGTCTACCTATGT 328
 /organism="Lotus japonicus"

Db 5008 TGAAATGAAAGGGATCACTAATATATTAGTTATGTTGATGATATTCTATGGT 5067
 /organism="Lotus japonicus"

Db 327 TAAAGGTTGATGGTATACGTGTTATCTCTATGTTGATGCTATGTTAC 268
 /organism="Lotus japonicus"

QY 5068 TGAAGTTCAACAAAGTATTGTAATTGTGCTGATCATTGAGAGCATTTGAGT 5222
 /organism="Lotus japonicus"

Db 267 AGGAAGTGGAGATGTTGACAGTTAACAGGATGTTGAAT 208
 /organism="Lotus japonicus"

QY 5128 TAAGTGTGTTGAAATATCAATTATCTGTGTTGATTGTTGAAACCGATCTGG 518
 /organism="Lotus japonicus"

Db 207 GACCGATCTTGGAAAGATGACATCTTCTGTTGATGCGAGCAAAGCACATGA 148
 /organism="Lotus japonicus"

QY 5188 TTATTTATTCAGAALATTCTCAGAAMATTACTAGGATTCAACTAGATGA 5247
 /organism="Lotus japonicus"

Db 147 AATATTGCTGTTGAGAAAGTATGCAAAAGGAGTCCTCAAGAGTCACTGAAGA 88
 /organism="Lotus japonicus"

QY 5248 CTCATATGGAAACATACCTGGATTGAGTACAA 5286
 /organism="Lotus japonicus"

Db 87 GTGCCATCCACAGTACTCCATGATCAAGGAAA 49
 /organism="Lotus japonicus"

RESULT 15

AG264218/C

LOCUS AG264218 490 bp DNA linear GSS 22-JUL-2003
 DEFINITION Lotus corniculatus var. japonicus DNA, clone:LJ59017_sfi, genomic
 survey sequence.

ACCESSION AG264218

VERSION 1.1 GI:26664057

KEYWORDS GSS.

SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotteae; Lotus.

REFERENCE 1

AUTHORS Sato, S., Nakamura, Y. and Tabata, S.

TITLE Lotus japonicus TAC End sequences

JOURNAL Published Only in Database (2002)

REFERENCE 2

AUTHORS Sato, S.

TITLE Direct Submission

JOURNAL The First Laboratory for Plant Gene Research; 2-6-7
 Kazusa-hamatai, Kisarazu, Chiba 292-0818, Japan
 (E-mail: ssato@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/,
 Tel:81-438-52-3935(ex.2336), Fax:81-438-52-3934)
 Location/Qualifiers
 1. 490
 /organism="Lotus corniculatus var. japonicus"

FEATURES source

Job time : 9830 secs

Search completed: April 11, 2004, 21:13:15

This Page Blank (uspto)